THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Stochastic PDEs with extremal properties

Máté Gerencsér

Doctor of Philosophy
University of Edinburgh
2016

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Máté Gerencsér*)

Acknowledgements

The author is deeply grateful for the guidance of his supervisor, Professor István Gyöngy, whose help cannot be overestimated. Were it not for his wisdom, patience, and encouragement, this work could not have been done.

The author also thanks the support of the Principal's Career Development Scholarship as well as the conference supports by the School of Mathematics, the Laura Wisewell Fund, the IMU Itô Travel Award, and the Centro di Ricerca Matematica Ennio De Giorgi.

Abstract

We consider linear and semilinear stochastic partial differential equations that in some sense can be viewed as being at the "endpoints" of the classical variational theory by Krylov and Rozovskii [25]. In terms of regularity of the coefficients, the minimal assumption is boundedness and measurability, and a unique L_2-valued solution is then readily available. We investigate its further properties, such as higher order integrability, boundedness, and continuity. The other class of equations considered here are the ones whose leading operators do not satisfy the strong coercivity condition, but only a degenerate version of it, and therefore are not covered by the classical theory. We derive solvability in W_p^m spaces and also discuss their numerical approximation through finite different schemes.

Keywords— Stochastic PDEs, Cauchy problem, Moser's iteration, Harnack inequality, degenerate parabolicity, symmetric hyperbolic systems, finite differences, localization error

Lay Summary

In this thesis we investigate stochastic partial differential equations. These equations describe the evolution of a random quantity in time. They are infinite dimensional in that at every instance of time the quantity consists of infinitely many values. An illustrative example is the heat equation describing the propagation of temperature in a certain medium: at any given time the state of the system is given by the collection of values of temperature at each point in space, that is, a function of space. If the source of the heat is random or there are other uncertainties in the system then the equation modelling the evolution will have stochastic terms.

For any mathematical model it is crucial that the model itself is self-consistent, that is, that the equation has a solution in a reasonably defined sense. When the solutions exist, one might be interested in further properties of it and study whether it is a bounded function, a smooth one, or whether it is possible to approximate it in a reasonable manner. The practical motivations of the latter is also quite clear: while the existence of solutions may be provable in large generality, they are rarely available explicitly, and therefore one would like to have methods that are easily implementable numerically and yield functions that are close to the true solution.

On the other hand, it is desirable that such properties do not require too much from the equation, limiting the range of applicability. This is one of the motivations to study equations which are not, or are barely covered by the usual methods but may very well naturally appear in applications. Studying solvability, regularity, and numerics of such equations we extend (and in some cases, sharpen) the known results to a wider class of equations.

Contents

Chapter 1

Introduction

Stochastic partial differential equations (SPDEs) have been the subject of very active research in the past decades, motivated by a wide variety of applications. One of the main approaches to analyse these equation, also referred to as the "variational approach", was developed in [34] and [25]. Following the latter reference, the main well-posedness result can be formulated in an abstract setting as follows.

Fix a terminal time $T > 0$, and consider a probability space $(\Omega, \mathcal{F}, P)$, equipped with a complete, right-continuous filtration $(\mathcal{F}_t)_{t\in[0,T]}$, and let $(w^k)_{k=1}^\infty$ be a sequence of independent $(\mathcal{F}_t)_{t\in[0,T]}$-Wiener martingales. The predictable σ-algebra on $\Omega\times[0,T]$ is denoted by $\mathcal{P}$. Let H be a separable real Hilbert space and let V be a separable, reflexive, real Banach space continuously and densely embedded in H. Identifying H with its dual, this induces the continuous and dense inclusions $V \hookrightarrow H \hookrightarrow V^*$, with the identity $\langle v, h\rangle = (v, h)$ for $v \in V, h \in H$, where $\langle\cdot,\cdot\rangle$ is the duality pairing and $(\cdot,\cdot)$ is the inner product in H. This is also referred to as a Gelfand triple. An important example, and the one most relevant for the present work, is the triple $H^s \hookrightarrow H^{s-1} \hookrightarrow H^{s-2}$ for some $s \in \mathbb{R}$, where $H^s = W^s_2$ are Sobolev spaces, introduced in detail below. Consider the stochastic evolution equation

$$u_t = u_0 + \int_0^t A_s(u_s)\,ds + \sum_{k=1}^\infty \int_0^t B^k_s(u_s)\,dw^k_s, \tag{1.0.1}$$

for $t \in [0, T]$, under the following assumptions (note that whenever it does not cause confusion, certain arguments are suppressed, as, for example, ω is in (1.0.1)).

Assumption 1.0.1. The operators A and $B = (B^k)_{k=1}^\infty$ are $\mathcal{P}\times\mathcal{B}(V)$-measurable functions from $\Omega \times [0, T] \times V$ to V^* and $l_2(H)$, respectively, such that for all

$u, v, w \in V$, $\omega, t \in \Omega \times [0, T]$:

(i) (Monotonicity)

$$2\langle u - v, A(u) - A(v)\rangle + \sum_{k=1}^{\infty} |B^k(u) - B^k(v)|_H^2 \le K|u - v|_H^2;$$

(ii) (Strong coercivity)

$$2\langle u, A(u)\rangle + \sum_{k=1}^{\infty} |B^k(u)|_H^2 \le -\lambda\|u\|_V^2 + K|u|_H^2 + f;$$

(iii) (Linear growth)

$$\|A(u)\|_{V^*}^2 \le K\|u\|_V^2 + f, \qquad \sum_{k=1}^{\infty} |B^k(u)|_H^2 \le \|u\|_V^2 + f;$$

(iv) (Hemicontinuity)

$$\lim_{\varepsilon \to 0}(u, A(v + \varepsilon w)) = (u, A(v))$$

.

Here $K \ge 0$ and $\lambda > 0$ are constants, while f is an adapted nonnegative process such that

$$E \int_0^T f_s \, ds < \infty.$$

The initial condition u_0 is assumed to be an $\mathcal{F}_0$-measurable H-valued random variable. The solution of (1.0.1) is then understood as follows.

Definition 1.0.1. An $H-$valued adapted continuous process $(u_t)_{t\in[0,T]}$ is called a solution of (1.0.1) if $u_t \in V$ for almost all $\omega, t \in \Omega \times [0, T]$, and almost surely

$$\int_0^T \|u_t\|_V^2 dt < \infty$$

and

$$(u_t, v) = (u_0, v) + \int_0^t \langle A_s(u_s), v\rangle \, ds + \sum_{k=1}^{\infty} \int_0^t (B_s^k(u_s), v) dw_s^k$$

for all $t \in [0, T]$ and $v \in V$.

Theorem 1.0.1. *Let Assumption 1.0.1 hold. Then* (1.0.1) *admits a unique (up to indistinguishability) solution* $(u_t)_{t\in[0,T]}$*, and moreover, there exists a constant*

$N = N(\lambda, K, T)$, *such that*

$$E \sup_{t\in[0,T]} |u_t|_H^2 + E\int_0^T \|u_s\|_V^2\,ds \le N\left(E|u_0|_H^2 + E\int_0^T f_s\,ds\right).$$

Applying the theorem to the example $H^s \hookrightarrow H^{s-1} \hookrightarrow H^{s-2}$, one can get well-posedness results for the second order parabolic stochastic PDE

$$du = (I_{div} D_i(a^{ij}D_j u) + I_{nondiv}a^{ij}D_iD_ju + f(u, Du))\,dt + \sum_{k=1}^{\infty}(\sigma^{ik}D_iu + g^k(u))dw_t^k$$

in either divergence or nondivergence form (i.e. exactly one of *div* and *nondiv* is "true"), for sufficiently nice f and g. In terms of the coefficients a and σ, Assumption 1.0.1 here translates to

(i) A stochastic parabolicity condition:

$$(2a^{ij} - \sum_k \sigma^{ik}\sigma^{jk})_{i,j=1}^d \ge \lambda I$$

as symmetric matrices, for some $\lambda > 0$, where I is the identity matrix,

(ii) Certain smoothness assumptions in the spatial variable, depending on s and the form of the equation.

In (ii), the minimal smoothness requirement occurs when $s = 1$ and $div =$ true, in which case only boundedness is required from the coefficients. This is the topic of Chapter 2, where we investigate the further properties of the unique L_2-valued solution provided by Theorem 1.0.1. While the established properties of the solutions are available through much easier arguments in the case of more regular coefficients, assuming the minimal conditions not only provides more generality, but also sharper estimates. These results can also be used to derive new existence results for a wide class of semilinear equations.

The (excluded) endpoint of (i) is $\lambda = 0$, in this case solvability in H^s is proved in [26]. Such degeneracy may arise naturally from applications, particularly in the nonlinear filtering problem. It is also useful to have a theory that includes tha $\lambda = 0$ case for studying truncated equations, which may appear in numerical approximations. We discuss solvability in W_p^m in Chapter 3.

As explicit solutions are rarely available, disctretization of SPDEs are of great interest. While the literature is extensive, similarly to the theoretical results, much fewer is known for degenerate equations. This is what we investigate in

Chapter 4, focusing on the acceleration of the rate of convergence of finite difference approximations and the error of localization.

1.1 Notations

The probabilistic setup is already introduced above, for other basic notions in stochastic analysis used in the following such as stopping times, stochastic integration, continuous martingales and their quadratic variation process we refer to [16] or [35]. For a fixed $d \geq 1$, we denote $B_R = \{x \in \mathbb{R}^d : |x| < R\}$ for $R \geq 0$. The Lebesque measure of a set A is denoted by $|A|$. For a domain $A \subset \mathbb{R}^d$, $p \in (0, \infty]$, and a Hilbert space H the norm in $L_p(A, H)$ is denoted by $|\cdot|_{L_p}$ or $|\cdot|_p$, while the norm in $L_p([s,r] \times A, H)$ is denoted by $\|\cdot\|_{p,[s,r]\times A}$, or, whenever omitting the domain does not cause confusion, by $\|\cdot\|_p$. Similarly, the norm in $L_p([s,r], L_q(A,H))$ is denoted by $\|\cdot\|_{p,q}$. The target space H will usually be omitted, as it will be clear from the context which function takes values where. For a nonnegative integer m, $W_p^m = W_p^m(\mathbb{R}^n)$ denotes the Sobolev space consisting of functions such that their distributional derivatives up to order m are in L_p. Here and in the following when we talk about "derivatives up to order m", we understand the inclusion of the zero-th derivative, that is, the function itself. When $p = 2$, we often use the notation $W_2^m = H^m$. The space of smooth functions compactly supported on a domain $A \subset \mathbb{R}^d$ is denoted by $C_c^\infty(A)$. The closure of $C_c^\infty(A)$ in the H^1 norm is denoted by $H_0^1(A)$, and its dual by H^{-1}. For (distributional) derivatives of functions on $\mathbb{R}^d$ we use the notations

$$D_i = \partial_i = \frac{\partial}{\partial x^i}, \quad D_{ij} = D_i D_j, \quad \partial_v = \sum_{i=1}^d v^i D_i, \quad \nabla = D = (D_1, \ldots, D_d)$$

for $i, j = 1, \ldots d$. For a multiindex $\alpha = (\alpha_1, \ldots, \alpha_d) \in \{0, 1, \ldots\}^d$, we define its length $|\alpha| = \sum_i \alpha_i$ and $D^\alpha = D_1^{\alpha_1} \cdots D_d^{\alpha_d}$. By $\inf, \sup$, etc. we always mean essential ones, although this often (for example in Theorem 1.0.1 above) will agree with the true $\inf, \sup$, etc. Indicators of a set A is denoted by $\mathbf{1}_A$. In the following the summation convention with respect to repeated indices is used whenever not indicated otherwise. Constants in the calculations, usually denoted by C or N, may change from line to line, but their dependency is indicated in the relevant statement.

1.2 Preliminaries

We will often use classical inequalities such as Hölder's, Young's, Burkholder-Davis-Gundy, Doob's, etc. Below, let us collect some useful but perhaps less well-known technical lemmas from the theory of function spaces and stochastic analysis, respectively, that will also be used on one or more occasions.

Lemma 1.2.1 ((II.3.4),[29])**.** *Let $Q \subset \mathbb{R}^d$ be a Lipschitz domain and suppose that $v \in L_2([0,T], H_0^1(Q)) \cap L_\infty([0,T], L_2(Q))$. Let $r, q \in (2,\infty)$, satisfying $1/r + d/2q = d/4$. Then v belongs to $L_r([0,T], L_q(Q))$, and*

$$\left(\int_0^T \left(\int_Q |v_t|^q dx\right)^{r/q} dt\right)^{2/r} \leq N \left(\sup_{0\leq t\leq T} \int_Q |v_t|^2 dx + \int_0^T \int_Q |\nabla v_t|^2 dx dt\right)$$

with $N = N(d, |Q|, T)$

Lemma 1.2.2 ((II.5.4),[29])**.** *Let $\rho > 0$ and $v \in H^1(B_\rho)$ such that on $A \subset B_\rho$, $v = 0$. Then*

$$\int_{B_\rho} v^2 \, dx \leq N \frac{\rho^{2(d+1)}}{|A|^2} \int_{B_\rho} |\nabla v|^2 \, dx,$$

with $N = N(d)$.

Lemma 1.2.3 (IV.4.7/IV.4.31,[35])**.** *Let X be a non-negative adapted right-continuous process, and let A be a non-decreasing continuous process such that*

$$E(X_\tau | \mathcal{F}_0) \leq E(A_\tau | \mathcal{F}_0)$$

for any bounded stopping time τ. Then for any $x, y > 0$,

$$P(\sup_{t\leq T} X_t \geq x, A_T \leq y) \leq y/x,$$

and for any $\sigma \in (0,1)$

$$E \sup_{t\leq T} X_t^\sigma \leq \sigma^{-\sigma}(1-\sigma)^{-1} E A_T^\sigma.$$

Lemma 1.2.4 ([33])**.** *Let $a = (a^{ij}(x))$ be a function defined on $\mathbb{R}^d$, with values in the set of non-negative $m \times m$ matrices, such that a and its derivatives in x up second order are bounded in magnitude by a constant K. Let V be a symmetric $m \times m$ matrix. Then*

$$|Da^{ij}V^{ij}|^2 \leq N a^{ij} V^{ik} V^{jk}$$

for every $x \in \mathbb{R}^d$, where $N = N(K, d)$.

Lemma 1.2.5 ([12]). *Let $y = (y_t)_{t\in[0,T]}$ and $F = (F_t)_{t\in[0,T]}$ be adapted nonnegative stochastic processes and let $m = (m_t)_{t\in[0,T]}$ be a continuous local martingale such that*

$$dy_t \le (Ny_t + F_t)\,dt + dm_t \quad \text{on } [0,T] \tag{1.2.2}$$

$$d\langle m\rangle_t \le (Ny_t^2 + y_t^{2(1-\rho)}G_t^{2\rho})\,dt \quad \text{on } [0,T], \tag{1.2.3}$$

with some constants $N \ge 0$ and $\rho \in [0,1/2]$, and a nonnegative adapted stochastic process $G = (G_t)_{t\in[0,T]}$, such that

$$\int_0^T G_t\,dt < \infty\,(a.s.),$$

where $\langle m\rangle$ is the quadratic variation process for m. Then for any $q > 0$

$$E\sup_{t\le T} y_t^q \le CEy_0^q + CE\left\{\int_0^T (F_t + G_t)\,dt\right\}^q$$

with a constant $C = C(N,q,\rho,T)$.

Chapter 2

Discontinuous coefficients

In this chapter we investigate (1.0.1) in divergence form, with bounded but possibly discontinuous coefficients. The general theory covers this case, with the triple $H^1 \hookrightarrow L_2 \hookrightarrow H^{-1}$, and therefore one knows the existence of an L_2-valued (for almost all ω, t, H^1-valued) solution. Deterministic theory suggests, however, that more can be said about the solution: [6], [32], and [31] established Hölder-continuity of the solutions of elliptic equations $Lu = 0$, with merely bounded, measurable, and elliptic coefficients. This is the celebrated De Giorgi-Nash-Moser theory, which turned out to be a key result in the theory of nonlinear PDEs. It is a natural question to ask whether such results hold for SPDEs. This was investigated in the author's collaboration with Konstantinos Dareiotis in the papers [3], [4]. The content of this chapter is based on this work.

Remark 2.0.1. The main purpose is to tackle the problems arising due to the stochastic nature of the equation, and therefore we did not attempt full generality. The directions towards which generalizations are available and are relatively straightforward include unbounded lower order coefficients ([17],[29]), different integrability exponents in space and time ([3],[5],[29]), and semilinear equations ([29]), with the nonlinear term growing slightly superlinearly. In fact, as seen in [5], some nonlinearities of the leading order can also be included. Given that the estimates for treating the additional terms arising in these generalizations can be found in [29], we do not include (let alone unify) these approaches.

2.1 Global supremum estimates

Consider the equation

$$du_t = (L_t u_t + \partial_i f_t^i + f_t^0)dt + (M_t^k u_t + g_t^k)dw_t^k, \; u_0 = \psi, \tag{2.1.1}$$

on a bounded Lipschitz domain Q with 0 boundary condition, where

$$L_t u = \partial_j(a^{ij}_t \partial_i u) + b^i_t \partial_i u + c_t u, \ \ M^k_t u = \sigma^{ik}_t \partial_i u_t + \mu^k_t u.$$

We aim to derive global (i.e. up to the space-time boundary) estimates for the supremum norm of the solution. Solutions are understood via Definition 1.0.1, on the triple $H^1_0(Q) \hookrightarrow L_2(Q) \hookrightarrow H^{-1}(Q)$. We also get the existence and uniqueness of the solution by Theorem 1.0.1, under the following assumptions.

Assumption 2.1.1. i) The coefficients a^{ij}, b^i and c are real-valued $\mathcal{P} \times \mathcal{B}(Q)$ measurable functions on $\Omega \times [0,T] \times Q$ and are bounded by a constant $K \geq 0$, for any $i, j = 1, ..., d$. The coefficients $\sigma^i = (\sigma^{ik})^\infty_{k=1}$ and $\mu = (\mu^k)^\infty_{k=1}$ are l_2-valued $\mathcal{P} \times Q$-measurable functions on $\Omega \times [0,T] \times Q$ such that

$$\sum_i \sum_k |\sigma^{ik}_t(x)|^2 + \sum_k |\mu^k_t(x)|^2 \leq K \quad \text{for all } \omega, \, t \text{ and } x,$$

ii)f^l, for $l \in \{0, ..., d\}$, and $g = (g^k)^\infty_{k=1}$ are $\mathcal{P} \times \mathcal{B}(Q)$-measurable functions on $\Omega \times [0,T] \times Q$ with values in $\mathbb{R}$ and l_2, respectively, such that

$$E(\sum_{l=0}^d \|f^l\|_2^2 + \||g|_{l_2}\|_2^2) < \infty$$

iii) ψ is an $\mathcal{F}_0$-measurable random variable in $L_2(Q)$ such that $E|\psi|_2^2 < \infty$

Assumption 2.1.2. There exists a constant $\lambda > 0$ such that for all ω, t, x and for all $\xi = (\xi_1, ...\xi_d) \in \mathbb{R}^d$ we have

$$a^{ij}_t(x)\xi_i\xi_j - \frac{1}{2}\sigma^{ik}_t(x)\sigma^{jk}_t(x)\xi_i\xi_j \geq \lambda|\xi|^2.$$

Let

$$\Gamma_d = \left\{ (r,q) \in (1,\infty]^2 \middle| \frac{1}{r} + \frac{d}{2q} < 1 \right\}.$$

The following is our main result on global boundedness.

Theorem 2.1.1. *Suppose that Assumptions 2.1.1 and 2.1.2 hold, and let u be the unique L_2-solution of equation* (2.1.1). *Then for any $(r,q) \in \Gamma_d$ and $\eta > 0$,*

$$E\|u\|^\eta_\infty \leq NE(|\psi|^\eta_\infty + \|f^0\|^\eta_{r,q} + \sum_{i=1}^d \|f^i\|^\eta_{2r,2q} + \||g|_{l_2}\|^\eta_{2r,2q}), \tag{2.1.2}$$

where $N = N(\eta, r, q, d, K, \lambda, |Q|, T)$.

We start by proving an Itô's formula for the p-th norm, from which we then derive some "energy inequality-like" estimates.

Lemma 2.1.2. *Suppose that u satisfies equation* (2.1.1), $f^l \in L_p(\Omega \times [0,T], \mathcal{P}; L_p(Q))$ *for* $l \in \{0, ..., d\}$, $g \in L_p(\Omega \times [0,T], \mathcal{P}; L_p(Q))$, *and* $\psi \in L_p(\Omega, \mathcal{F}_0; L_p(Q))$ *for some* $p \geq 2$. *Then there exists a constant* $N = N(d, K, \lambda, p)$, *such that*

$$E \sup_{t \leq T} |u_t|_p^p + E \int_0^T \int_Q |\nabla u_s|^2 |u_s|^{p-2} dx ds \leq N E(|\psi|_p^p + \sum_{l=0}^d \|f^l\|_p^p + \||g|_{l_2}\|_p^p). \tag{2.1.3}$$

Moreover, almost surely

$$\begin{aligned} \int_Q |u_t|^p dx = &\int_Q |u_0|^p dx + p \int_0^t \int_Q (\sigma_s^{ik} \partial_i u_s + \mu^k u_s + g^k) u_s |u_s|^{p-2} dx dw_s^k \\ &+ \int_0^t \int_Q -p(p-1) a_s^{ij} \partial_i u_s |u_s|^{p-2} \partial_j u_s - p(p-1) f_s^i \partial_i u_s |u_s|^{p-2} dx ds \\ &+ \int_0^t \int_Q p(b_s^i \partial_i u_s + c_s u_s + f_s^0) u_s |u_s|^{p-2} dx ds \\ &+ \frac{1}{2} p(p-1) \int_0^t \int_Q \sum_{k=1}^\infty |\sigma_s^{ik} \partial_i u_s + \mu^k u_s + g_s^k|^2 |u_s|^{p-2} dx ds, \end{aligned} \tag{2.1.4}$$

for any $t \leq T$.

Proof. Consider the functions

$$\phi_n(r) = \begin{cases} |r|^p & \text{if } |r| < n \\ n^{p-2} \frac{p(p-1)}{2} (|r| - n)^2 + p n^{p-1} (|r| - n) + n^p & \text{if } |r| \geq n. \end{cases}$$

Then one can see that ϕ_n are twice continuously differentiable, and satisfy

$$|\phi_n(x)| \leq N|x|^2, \ |\phi_n'(x)| \leq N|x|, \ |\phi_n''(x)| \leq N,$$

where N depends only on p and $n \in \mathbb{N}$. We also have that for any $r \in \mathbb{R}$, $\phi_n(r) \to |r|^p$, $\phi_n'(r) \to p|r|^{p-2} r$, $\phi_n''(r) \to p(p-1)|r|^{p-2}$, as $n \to \infty$, and

$$\phi_n(r) \leq N|r|^p, \ \phi_n'(r) \leq N|r|^{p-1}, \ \phi_n''(r) \leq N|r|^{p-2}, \tag{2.1.5}$$

where N depends only on p. Then for each $n \in \mathbb{N}$ we have almost surely

$$
\begin{aligned}
\int_Q \phi_n(u_t)dx = &\int_Q \phi_n(u_0)dx + \int_0^t \int_Q (\sigma_s^{ik}\partial_i u_s + \mu^k u_s + g^k)\phi_n'(u_s)dxdw_s^k \\
&+ \int_0^t \int_Q -a_s^{ij}\partial_i u_s \phi_n''(u_s)\partial_j u_s - f^i\phi_n''(u_s)\partial_i u_s dxds \\
&+ \int_0^t \int_Q b_s^i \partial_i u_s \phi_n'(u_s) + c_s u_s \phi_n'(u_s) + f_s^0 \phi_n'(u_s) dxds \\
&+ \frac{1}{2}\int_0^t \int_Q \sum_{k=1}^{\infty} |\sigma_s^{ik}\partial_i u_s + \mu^k u_s + g_s^k|^2 \phi_n''(u_s)dxds, \qquad (2.1.6)
\end{aligned}
$$

for any $t \in [0,T]$ (see for example, Section 3 in [19]). By Young's inequality, and the parabolicity condition we have for any $\varepsilon > 0$,

$$
\begin{aligned}
\int_Q \phi_n(u_t)dx \leq\ & m_t^{(n)} + \int_Q \phi_n(u_0)dx \\
&+ \int_0^t \int_Q (-\lambda|\nabla u_s|^2 + \varepsilon|\nabla u_s|^2 + N\sum_{i=1}^{d} |f_s^i|^2)\phi_n''(u_s)dxds \\
&+ \int_0^t \int_Q (\epsilon|\nabla u_s|^2 + N|u_s|^2 + N\sum_{k=1}^{\infty} |g_s^k|^2)\phi_n''(u_s)dxds \\
&+ \int_0^t \int_Q (b_s^i\partial_i u_s + c_s u_s + f_s^0)\phi_n'(u_s)dxds, \qquad (2.1.7)
\end{aligned}
$$

where $N = N(d, K, \epsilon)$, and $m_t^{(n)}$ is the martingale from (2.1.6). One can check that the following inequalities hold,

i) $|r\phi_n'(r)| \leq p\phi_n(r)$

ii) $|r^2\phi''(r)| \leq p(p-1)\phi_n(r)$

iii) $|\phi_n'(r)|^2 \leq 4p\ \phi_n''(r)\phi_n(r)$

iv) $[\phi_n''(r)]^{p/(p-2)} \leq [p(p-1)]^{p/(p-2)}\phi_n(r)$,

which combined with Young's inequality imply,

i) $\partial_i u_s \phi_n'(u_s) \leq \epsilon\phi_n''(u_s)|\partial_i u_s|^2 + N\phi_n(u_s)$

ii) $|u_s\phi_n'(u_s)| \leq p\phi_n(u_s)$

iii) $|f_s^0\phi_n'(u_s)| \leq |f_s^0||\phi_n''(u_s)|^{1/2}|\phi_n(u_s)|^{1/2} \leq N|f_s^0|^p + N\phi_n(u_s)$

iv) $|u_s|^2\phi_n''(u_s) \leq N\phi_n(u_s)$

v) $\sum_k |g^k_s|^2\phi''_n(u_s) \leq N\phi_n(u_s) + N\Big(\sum_k |g^k_s|^2\Big)^{p/2}$

vi) $\sum_{i=1}^d |f^i_s|^2\phi''_n(u_s) \leq N\phi_n(u_s) + N\sum_{i=1}^d |f^i_s|^p$,

where N depends only on p and ϵ.

By choosing ϵ sufficiently small, and taking expectations we obtain

$$E\int_Q \phi_n(u_t)dx + E\int_0^t\int_Q |\nabla u_s|^2\phi''_n(u_s)dxds \leq NE\mathcal{K}_t + N\int_0^t E\int_Q \phi_n(u_s)dxds,$$

where $N = N(d, p, K, \lambda)$ and

$$\mathcal{K}_t = |\psi|^p_p + \int_0^t \sum_{l=0}^d |f^l_s|^p_p + |g_s|^p_p ds.$$

By Gronwall's lemma we get

$$E\int_Q \phi_n(u_t)dx + E\int_0^t\int_Q |\nabla u_s|^2\phi''_n(u_s)dxds \leq N\mathbb{E}\mathcal{K}_t$$

for any $t \in [0, T]$, with $N = N(T, d, p, K, \lambda)$. Going back to (2.1.7), using the same estimates, and the above relation, by taking suprema up to T we have

$$E\sup_{t\leq T}\int_Q \phi_n(u_t)dx \leq NE\mathcal{K}_t + E\sup_{t\leq T}|m^{(n)}_t|.$$

$$\leq NE\mathcal{K}_T + NE\left(\int_0^T\sum_k\left(\int_Q |\sigma^{ik}\partial_i u_s + \mu^k u_s + g^k_s||\phi''_n(u_s)\phi_n(u_s)|^{1/2}dx\right)^2 ds\right)^{1/2}$$

$$\leq NE\mathcal{K}_T + NE\left(\int_0^T\int_Q (|\nabla u_s|^2 + |u_s|^2 + \sum_{k=1}^\infty |g^k_s|^2)\phi''_n(u_s)dx\int_Q \phi_n(u_s)dxds\right)^{1/2}$$

$$\leq NE\mathcal{K}_T + \frac{1}{2}E\sup_{t\leq T}\int_Q \phi_n(u_t)dx < \infty,$$

where $N = N(T, d, p, K, \lambda)$. Hence,

$$E\sup_{t\leq T}\int_Q \phi_n(u_t)dx + E\int_0^T\int_Q |\nabla u_s|^2\phi''_n(u_s)dxds \leq NE\mathcal{K}_T,$$

and by Fatou's lemma we get (2.1.3). For (2.1.4), we go back to (2.1.6), and by letting a subsequence $n(k) \to \infty$ and using the dominated convergence theorem, we see that each term converges to the corresponding one in (2.1.4) almost surely, for all $t \leq T$. This finishes the proof.

$\square$

Corollary 2.1.3. *Let $\gamma > 1$ and denote $\kappa = 4\gamma/(\gamma-1)$. Suppose furthermore that $r, r', q, q' \in (1, \infty)$, satisfying $1/r + 2/r' = 1$ and $1/q + 2/q' = 1$. Suppose that u satisfies the conditions of Lemma 2.1.2 for any $p \in \{2\gamma^n, n \in \mathbb{N}\}$. Then, for any $p \in \{2\gamma^n, n \in \mathbb{N}\}$, almost surely, for all $t \leq T$*

$$\int_Q |u_t|^p dx + \frac{p^2}{4} \int_0^t \int_Q |\nabla u_t|^2 |u_t|^{p-2} dx ds \leq N' m_t$$

$$+ N \left[|\psi|_p^p + p^\kappa \|u\|^p_{r'p/2, q'p/2} + p^{-p} (\|f^0\|^p_{r,q} + \sum_{i=1}^d \|f^i\|^p_{2r,2q} + \||g|_{l_2}\|^p_{2r,2q}) \right], \quad (2.1.8)$$

where m_t is the martingale from (2.1.4), *and N, N' are constants depending only on $K, d, T, \lambda, |Q|, r, q, \gamma$.*

Proof. By Lemma 2.1.2, the parabolicity condition, and Young's inequality we have

$$\int_Q |u_t|^p dx + \frac{p^2}{4} \int_0^t \int_Q |\nabla u_s|^2 |u_s|^{p-2} dx ds \leq N' m_t + N_1 \left(\int_Q |\psi|^p dx \right.$$

$$\left. + \int_0^t \left[\int_Q p^2 |u_s|^p + p|f^0_s||u_s|^{p-1} + p^2 \sum_{i=1}^d |f^i_s|^2 |u_s|^{p-2} + p^2 |g_s|^2_{l_2} |u_s|^{p-2} dx \right] ds \right).$$

Then by Hölder's inequality we have

$$\int_0^t \int_Q |f^0_s||u_s|^{p-1} dx ds \leq \|f^0\|_{r,q} \|u\|^{p-1}_{q'(p-1)/2, r'(p-1)/2},$$

and by Young's inequality we obtain

$$\begin{aligned} p\|f^0\|_{r,q} \|u\|^{p-1}_{q'(p-1)/2, r'(p-1)/2} &\leq p^{-p} \|f^0\|^p_{r,q} + p^\kappa \|u\|^p_{r'(p-1)/2, q'(p-1)/2} \\ &\leq p^{-p} \|f^0\|^p_{r,q} + N_2 p^\kappa \|u\|^p_{r'p/2, q'p/2}. \end{aligned}$$

Similarly, for $n \geq 1$,

$$\begin{aligned} p^2 \int_0^t \int_Q |f^i_s|^2 |u_s|^{p-2} dx ds &\leq p^2 \|f^i\|^2_{2r,2q} \|u\|^{p-2}_{r'(p-2)/2, q'(p-2)/2} \\ &\leq p^{-p} \|f^i\|^p_{2r,2q} + p^\kappa \|u\|^p_{r'(p-2)/2, q'(p-2)/2} \\ &\leq p^{-p} \|f^i\|^p_{2r,2q} + N_3 p^\kappa \|u\|^p_{r'p/2, q'p/2}. \end{aligned}$$

The same holds for g in place of f^i. The case $n = 0$ can be covered separately with another constant N_4, and then N can be chosen to be $\max\{N_1(N_2+N_3), N_4\}$. This finishes the proof.

$\square$

Lemma 2.1.4. *Suppose that u satisfies equation* (2.1.1)*, $f^l \in L_p(\Omega \times [0,T], \mathcal{P}; L_p(Q))$ for $l \in \{0,...,d\}$, $g \in L_p(\Omega \times [0,T], \mathcal{P}; L_p(Q))$, and $\psi \in L_p(\Omega, \mathcal{F}_0; L_p(Q))$ for some $p \geq 2$. Then for any $0 < \eta < p$, and for any $\epsilon > 0$,*

$$E\left(\sup_{t\leq T}|u_t|_p^p + \frac{p^2}{4}E\int_0^T\int_Q |\nabla u_s|^2|u_s|^{p-2}dxds\right)^{\eta/p}$$

$$\leq \epsilon E\|u\|_\infty^\eta + N(\epsilon,p)E\left[|\psi|_p^\eta + \|f^0\|_1^\eta + \sum_{i=1}^d \|f^i\|_2^\eta + \||g|_{l_2}\|_2^\eta\right]$$

where $N(\epsilon,p)$ is a constant depending only on $\epsilon, \eta, K, d, T, \lambda, |Q|$, and p.

Proof. As in the proof of corollary 2.1.3, for any $\mathcal{F}_0-$measurable set B, we have almost surely

$$I_B\int_Q |u_t|^p dx + \frac{p^2}{4}I_B\int_0^t\int_Q |\nabla u_s|^2|u_s|^{p-2}dxds \leq N' I_B m_t + N_1 I_B\Bigg(\int_Q |\psi|^p dx$$

$$+\int_0^t\left[\int_Q p^2|u_s|^p + p|f_s^0||u_s|^{p-1} + p^2\sum_{i=1}^d |f_s^i|^2|u_s|^{p-2} + p^2|g_s|_{l_2}^2|u_s|^{p-2}dx\right]ds\Bigg), \tag{2.1.9}$$

for any $t \in [0,T]$. The above relation, by virtue of Gronwal's lemma implies that for any stopping time $\tau \leq T$

$$\sup_{t\leq T} EI_B\int_Q |u_{t\wedge\tau}|^p dx + EI_B\int_0^\tau\int_Q |\nabla u_s|^2|u_s|^{p-2}dxds \leq NEI_B\mathscr{V}_\tau, \tag{2.1.10}$$

where

$$\mathscr{V}_t := \int_Q |\psi|^p dx + \int_0^t\int_Q |f_s^0||u_s|^{p-1} + \sum_{i=1}^d |f_s^i|^2|u_s|^{p-2} + |g_s|_{l_2}^2|u_s|^{p-2}dxds.$$

Going back to (2.1.9), and taking suprema up to τ and expectations, and having in mind (2.1.10), gives

$$E\sup_{t\leq\tau} I_B\int_Q |u_t|^p dx \leq NE\sup_{t\leq\tau} I_B|m_t| + NEI_B\mathscr{V}_\tau.$$

By the Burkholder-Gundy-Davis inequality and (2.1.10) we have

$$
\begin{aligned}
E\sup_{t\le\tau} I_B|m_t| &\le NEI_B\left(\int_0^\tau\left(\int_Q|u_t|^{p-2}\left(|\nabla u_t|+|u_t|+|g|_{l_2}\right)dx\right)^2 dt\right)^{1/2}\\
&\le NEI_B\left(\int_0^\tau\int_Q|u_t|^p dx\int_Q(|\nabla u_t|^2+|u_t|^2+|g|_{l_2}^2)|u|^{p-2}dxdt\right)^{1/2}\\
&\le \frac{1}{2}E\sup_{t\le\tau} I_B\int_Q|u_t|^p dx+NEI_B\mathscr{V}_\tau.
\end{aligned}
$$

Hence,

$$
E\sup_{t\le\tau} I_B\int_Q|u_t|^p dx\le NEI_B\mathscr{V}_\tau,
$$

which combined with (2.1.10), by virtue of Lemma 1.2.3 gives

$$
E\left(\sup_{t\le T}|u_t|_p^p+\frac{p^2}{4}E\int_0^T\int_Q|\nabla u_s|^2|u_s|^{p-2}dxds\right)^{\eta/p}\le NE\mathscr{V}_T^{\eta/p}
$$

$$
\begin{aligned}
&\le NE\left[|\psi|_p^p+\|u\|_\infty^{p-1}\|f^0\|_1+\|u\|_\infty^{p-2}\left(\sum_{i=1}^d\|f^i\|_2^2+\||g|_{l_2}\|_2^2\right)\right]^{\eta/p}\\
&\le \epsilon E\|u\|_\infty^\eta+NE\left[|\psi|_p^\eta+\|f^0\|_1^\eta+\sum_{i=1}^d\|f^i\|_2^\eta+\||g|_{l_2}\|_2^\eta\right],
\end{aligned}
$$

which brings the proof to an end. □

Proof of Theorem 2.1.1 Throughout the proof, the constants N in our calculations will be allowed to depend on η, r, q as well as on the structure constants. Notice that we may, and we will assume that $r, q<\infty$. Without loss of generality we assume that the right hand side in (2.1.2) is finite. Also, in the first part of the proof we make the assumption that ψ, f^l, $l=0,\dots,d$, and g are bounded by a constant M. In particular, by (2.1.3), $u\in L_\eta(\Omega, L_{r,q})$ for any η, r, q.

Let us introduce the notation

$$
\mathcal{M}_{r,q,p}(t)=\|\mathbf{1}_{[0,t]}f^0\|_{r,q}^p+\sum_{i=1}^d\|\mathbf{1}_{[0,t]}f^i\|_{2r,2q}^p+\|\mathbf{1}_{[0,t]}|g|_{l_2}\|_{2r,2q}^p.
$$

Since $(r,q)\in\Gamma_d$, if we define r' and q' by $1/r+2/r'=1$, $1/q+2/q'=1$, we have

$$
\frac{d}{4}<\frac{1}{r'}+\frac{d}{2q'}=:\gamma\frac{d}{4}
$$

for some $\gamma > 1$. Then $\hat{r} = \gamma r'$ and $\hat{q} = \gamma q'$ satisfy

$$\frac{1}{\hat{r}} + \frac{d}{2\hat{q}} = \frac{d}{4}.$$

By applying Lemma 1.2.1 to $\hat{r}, \hat{q}$, and $\bar{v} = |v|^{p/2}$, we have, for any $p \geq 2$

$$E\left[|\psi|_\infty^\eta \vee \left(\int_0^T \left(\int_Q |v_t|^{\hat{q}p/2} dx\right)^{\hat{r}/\hat{q}} dt\right)^{2\eta/\hat{r}p}\right]$$

$$\leq E\left[|\psi|_\infty^\eta \vee N^{\eta/p}\left(\sup_{0\leq t\leq T}\int_Q |v_t|^p dx + \frac{p^2}{4}\int_0^T\int_Q |\nabla v_t|^2 |v_t|^{p-2} dxdt\right)^{\eta/p}\right]. \tag{2.1.11}$$

To estimate the right-hand side above, first notice that, if $p = 2\gamma^n$ for some n, then by taking supremum in (2.1.8), we have for any stopping time $\tau \leq T$, and any $\mathcal{F}_0-$ measurable set B,

$$I_B \sup_{0\leq s\leq \tau}\int_Q |v_s|^p dx$$

$$\leq NI_B\left(|\psi|_\infty^p + p^\kappa \|\mathbf{1}_{[0,\tau]}v\|^p_{r'p/2,q'p/2} + p^{-p}\mathcal{M}_{r,q,p}(\tau)\right) + N' I_B \sup_{0\leq s\leq\tau}|m_s|, \tag{2.1.12}$$

By the Davis inequality we can write

$$EI_B \sup_{0\leq s\leq\tau}|m_s| \leq NEI_B\left(\int_0^\tau \sum_k\left(\int_Q p(\sigma_s^{ik}\partial_i v_s + \mu^k v_s + g^k)v_s|v_s|^{p-2}dx\right)^2 ds\right)^{\frac{1}{2}}$$

$$\leq NEI_B\left(\sup_{0\leq s\leq\tau}\int_Q |v_s|^p dx\right)^{\frac{1}{2}}\left(\int_0^\tau\int_Q p^2\sum_k |\sigma_s^{ik}\partial_i v_s + \mu^k v_s + g^k|^2 |v_s|^{p-2} dxds\right)^{\frac{1}{2}}$$

Applying Young's inequality and recalling the already seen estimates in the proof of Corollary 2.1.3 (i) for the second term yields

$$EI_B \sup_{0\leq s\leq\tau}|m_s| \leq \varepsilon EI_B\left(\sup_{0\leq s\leq\tau}\int_Q |v_s|^p dx\right.$$

$$\left.+\frac{N}{\varepsilon}p^2\int_0^\tau\int_Q |\nabla v_s|^2|v_s|^{p-2}dxds + p^\kappa\|\mathbf{1}_{[0,\tau]}v\|^p_{r'p/2,q'p/2} + p^{-p}\|\mathbf{1}_{[0,\tau]}|g|_{l_2}\|^p_{2r,2q}\right)$$

for any $\varepsilon > 0$. With the appropriate choice of ε, combining this with (2.1.12) and using (2.1.8) once again, now without taking supremum, we get

$$EI_B\left(\sup_{0\le s\le\tau}\int_Q |v_s|^p dx+\frac{p^2}{4}\int_0^\tau\int_Q |\nabla v_s|^2|v_s|^{p-2}dxds\right)$$

$$\le NEI_B\Big(|\psi|_\infty^p+p^2\int_0^\tau\int_Q |\nabla v_s|^2|v_s|^{p-2}dxds+p^\kappa\|\mathbf{1}_{[0,\tau]}v\|^p_{r'p/2,q'p/2}+p^{-p}\mathcal{M}_{r,q,p}(\tau)\Big)$$
$$\le NEI_B\left(|\psi|_\infty^p+p^\kappa\|\mathbf{1}_{[0,\tau]}v\|^p_{r'p/2,q'p/2}+p^{-p}\mathcal{M}_{r,q,p}(\tau)\right)+N'EI_Bm_\tau,$$

and the last expectation vanishes. Now consider

$$X_t=|\psi|_\infty^p\vee\left(\sup_{0\le s\le t}\int_Q |v_s|^p dx+\frac{p^2}{4}\int_0^t\int_Q |\nabla v_s|^2|v_s|^{p-2}dxds\right)$$

and

$$A_t=Cp^\kappa\left(|\psi|_\infty^p\vee\|\mathbf{1}_{[0,t]}v\|^p_{r'p/2,q'p/2}+p^{-p}\mathcal{M}_{r,q,p}(t)\right)$$

for a large enough, but fixed C. The argument above gives that

$$\begin{aligned}EI_BX_\tau&\le EI_B\left(|\psi|_\infty^p+\sup_{0\le s\le\tau}\int_Q |v_s|^p dx+\frac{p^2}{4}\int_0^\tau\int_Q |\nabla v_s|^2|v_s|^{p-2}dxds\right)\\&\le NEI_B\left(|\psi|_\infty^p+p^\kappa\|\mathbf{1}_{[0,\tau]}v\|^p_{r'p/2,q'p/2}+p^{-p}\mathcal{M}_{r,q,p}(\tau)\right)\le EI_BA_\tau.\end{aligned}$$

Therefore the condition of Proposition 1.2.3 is satisfied, and thus for $\eta<p$ we obtain

$$E\left(|\psi|_\infty^p\vee\left(\sup_{0\le t\le T}\int_Q |v_t|^p dx+\frac{p^2}{4}\int_0^T\int_Q |\nabla v_t|^2|v_t|^{p-2}dxdt\right)\right)^{\eta/p}$$

$$\begin{aligned}&\le (Np^{\kappa+1})^{\eta/p}\frac{p}{p-\eta}E\left(|\psi|_\infty^p\vee\|v\|^p_{r'p/2,q'p/2}+p^{-p}\mathcal{M}_{r,q,p}(T)\right)^{\eta/p}\\&\le (Np^{\kappa+1})^{\eta/p}\frac{p}{p-\eta}E\left(|\psi|_\infty^\eta\vee\|v\|^\eta_{r'p/2,q'p/2}+p^{-\eta}\mathcal{M}_{r,q,\eta}(T)\right).\end{aligned}\tag{2.1.13}$$

Let us choose $p=p_n=2\gamma^n$ for $n\ge 0$, and use the notation $c_n=(Np_n^{\kappa+1})^{\eta/p_n}\frac{p_n}{p_n-\eta}$. Upon combining (2.1.11) and (2.1.13), for $p_n>\eta$ we can write the following inequality, reminiscent of Moser's iteration:

$$E|\psi|_\infty^\eta\vee\|v\|^\eta_{r'p_{n+1}/2,q'p_{n+1}/2}\le c_nE\left[|\psi|_\infty^\eta\vee\|v\|^\eta_{r'p_n/2,q'p_n/2}+Np_n^{-\eta}\mathcal{M}_{r,q,\eta}(T)\right].\tag{2.1.14}$$

Consider the minimal $n_0=n_0(d,\eta)$ such that $p_{n_0}>2\eta$. Taking any integer

$m \geq n_0$ we have

$$\prod_{n=n_0}^{m} c_n \leq \prod_{n=n_0}^{m} (N\gamma^{\kappa+1})^{\eta n/2\gamma^n} e^{2\eta/2\gamma^n}$$
$$= \exp\left[\log(N\gamma^{\kappa+1}) \sum_{n=n_0}^{m} \frac{\eta n}{2\gamma^n} + \sum_{n=n_0}^{m} \frac{\eta}{\gamma^n}\right] \leq N_0,$$

where N_0 does not depend on m. Also,

$$N \sum_{n=n_0}^{m} p_n^{-\eta} \leq N_1,$$

where N_1 does not depend on m. Therefore, by iterating (2.1.14) we get

$$\liminf_{m\to\infty} E|\psi|_\infty^\eta \vee \|v\|_{r'p_m/2,q'p_m/2}^\eta \leq N_0 N_1 E\mathcal{M}_{r,q,\eta}(T) + N_0 E|\psi|_\infty^\eta \vee \|v\|_{r'(p_{n_0}+1)/2,q'(p_{n_0}+1)/2}^\eta,$$

and thus by Fatou's lemma

$$E\|v\|_\infty^\eta \leq NE(|\psi|_\infty^\eta \vee \|v\|_{r'(p_{n_0}+1)/2,q'(p_{n_0}+1)/2}^\eta + \mathcal{M}_{r,q,\eta}(T)), \tag{2.1.15}$$

in particular, the left-hand side is finite.

By Lemma 2.1.4 we get

$$E\left(|\psi|_\infty^p \vee \left(\sup_{0\leq t\leq T} \int_Q |v_t|^p dx + \frac{p^2}{4}\int_0^T \int_Q |\nabla v_t|^2 |v_t|^{p-2} dxdt\right)\right)^{\eta/p}$$
$$\leq \epsilon E\|v\|_\infty^\eta + N(\epsilon, p)E\left(|\psi|_\infty^\eta + \mathcal{M}_{1,1,\eta}(T)\right) \tag{2.1.16}$$

for any $\epsilon > 0$. Combining (2.1.11) and (2.1.16) for $p = p_{n_0}$ gives

$$E|\psi|_\infty^\eta \vee \|v\|_{r'(p_{n_0}+1)/2,q'(p_{n_0}+1)/2}^\eta = E|\psi|_\infty^\eta \vee \|v\|_{\hat{r}p_{n_0}/2,q'p_{n_0}/2}^\eta$$
$$\leq \epsilon E\|v\|_\infty^\eta + N(\epsilon, p_{n_0})E\left(|\psi|_\infty^\eta + \mathcal{M}_{1,1,\eta}(T)\right). \tag{2.1.17}$$

Choosing ϵ sufficiently small, plugging (2.1.17) into (2.1.15), and rearranging yields the desired inequality

$$E\|v\|_\infty^\eta \leq NE(|\psi|_\infty^\eta + \mathcal{M}_{r,q,\eta}(T)). \tag{2.1.18}$$

As for the general case, set

$$\psi^{(n)} = \psi \wedge n \vee -n, \quad f^{l,(n)} = f^l \wedge n \vee -n, \quad g^{k,(n)} = g^k \wedge (n/k) \vee -(n/k),$$

define $\mathcal{M}^{(n)}_{r,q,p}$ correspondingly, and let v^n be the solution of the corresponding equation. This new data is now bounded by a constant, so the previous argument applies, and thus

$$E\|v^n\|^\eta_\infty \le NE(|\psi^{(n)}|^\eta_\infty + \mathcal{M}^{(n)}_{r,q,\eta}(T) \le NE(|\psi|^\eta_\infty + \mathcal{M}_{r,q,\eta}(T)).$$

Since $v^n \to v$ in $L_2(\Omega \times [0,T] \times Q)$, for a subsequence $k(n)$, $v^{k(n)} \to v$ for almost every ω, t, x. In particular, almost surely $\|v\|_\infty \le \liminf_{n\to\infty} \|v^{k(n)}\|_\infty$, and by Fatou's lemma

$$E\|v\|^\eta_\infty \le \liminf_{n\to\infty} E\|v^{k(n)}\|^\eta_\infty \le NE(|\psi|^\eta_\infty + \mathcal{M}_{r,q,\eta}(T)).$$

□

2.2 Semilinear SPDEs without growth condition

In this section, we will use the uniform norm estimates obtained in the previous section, to construct solutions for the following equation

$$du_t = (L_t u_t + f_t(u_t))dt + (M^k_t u_t + g^k_t)dw^k_t, u_0 = \psi \tag{2.2.19}$$

for $(t,x) \in [0,T] \times Q$, where f is a real function defined on $\Omega \times [0,T] \times Q \times \mathbb{R}$ and is $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d) \times \mathcal{B}(\mathbb{R})-$measurable.

Assumption 2.2.1. The function f satisfies the following

i) for all $r, r' \in \mathbb{R}$ and for all (ω, t, x) we have

$$(r - r')(f_t(x,r) - f_t(x,r')) \le K|r - r'|^2$$

ii) For all (ω, t, x), $f_t(x,r)$ is continuous in r

iii) for all $N > 0$, there exists a function $h^N \in L_2(\Omega \times [0,T] \times Q)$ with $E\|h^N\|_\infty < \infty$, such that for any (ω, t, x)

$$|f_t(x,r)| \le |h^N_t(x)|,$$

whenever $|r| \le N$.

iv) $E|\psi|_\infty + E\||g|_{l_2}\|_\infty < \infty$

Notice that other than the monotonicity and continuity, no polynomial (or any kind of) growth is assumed for f. Therefore even for the definition of the solution to make sense, u is required to be in L_∞:

Definition 2.2.1. A solution of equation (2.2.19) is an $\mathcal{F}_t-$adapted, strongly continuous process $(u_t)_{t\in[0,T]}$ with values in $L_2(Q)$ such that

i) $u_t \in H_0^1$, for $dP \times dt$ almost every $(\omega, t) \in \Omega \times [0,T]$

ii) $\int_0^T |u_t|_2^2 + |\nabla u_t|_2^2 dt < \infty$ (a.s.)

iii) almost surely, u is essentially bounded in (t,x)

iv) for all $\phi \in C_c^\infty(Q)$ we have with probability one

$$(u_t, \phi) = (\psi, \phi) + \int_0^t -(a_s^{ij}\partial_i u_s, \partial_j \phi) + (b_s^i \partial_i u_s + c_s u_s, \phi) + (f_s(u_s), \phi) ds$$

$$+ \int_0^t (M_s^k u_s + g_s^k, \phi) dw_s^k,$$

for all $t \in [0,T]$.

Notice that by Assumption 2.2.1 iii), and (iii) from Definition 2.2.1, the term $\int_0^t (f_s(u_s), \phi) ds$ is meaningful.

Theorem 2.2.1. *Under Assumptions 2.1.1, 2.1.2, and 2.2.1, there exists a unique solution of equation (2.2.19).*

Remark 2.2.1. From now on we can and we will assume that the function f is decreasing in r or else, by virtue of Assumption 2.2.1, we can replace $f_t(x,r)$ by $\tilde{f}_t(x,r) := f_t(x,r) - Kr$ and $c_t(x)$ with $\tilde{c}_t(x) := c_t(x) + K$.

We will need the following particular case from [2]. We consider two equations

$$du_t^i = (L_t u_t^i + f_t^i(u_t^i))dt + (M_t^k u_t^i + g_t^k) dw_t^k, \ u_0^i = \psi^i, \tag{2.2.20}$$

for $i = 1, 2$.

Assumption 2.2.2. The functions f^i, $i = 1, 2$, are appropriately measurable, and there exists $h \in L_2(\Omega \times [0,T] \times Q)$ and a constant $C > 0$, such that for any ω, t, x, and for any $r \in \mathbb{R}$ we have

$$|f_t^1(x,r)|^2 + |f_t^2(x,r)|^2 \leq C|r|^2 + |h_t(x)|^2.$$

Theorem 2.2.2. *Suppose that Assumptions 2.1.1, 2.1.2, and 2.2.2 hold. Let* u^i, $i = 1, 2$ *be the* L_2- *solutions of the equations in* (2.2.20), *for* $i = 1, 2$ *respectively. Suppose that* $f^1 \leq f^2$, $\psi^1 \leq \psi^2$ *and assume that either* f^1 *or* f^2 *satisfy Assumption 2.2.1. Then, almost surely and for any* $t \in [0, T]$, $u_t^1 \leq u_t^2$ *for almost every* $x \in Q$.

Proof of Theorem 2.2.1. We truncate the function f by setting

$$f_t^{n,m}(x,r) = \begin{cases} f_t(x,m) & \text{if } r > m \\ f_t(x,r) & \text{if } -n \leq r \leq m \\ f_t(x,-n) & \text{if } r < -n, \end{cases}$$

for $n,\ m \in \mathbb{N}$ we consider the equation

$$\begin{aligned} du_t^{n,m} &= (L_t u_t^{n,m} + f_t^{n,m}(u_t^{n,m}))dt + (M_t^k u_t^{n,m} + g_t^k)dw_t^k, \\ u_0^{n,m} &= \psi \end{aligned} \tag{2.2.21}$$

We first fix $m \in \mathbb{N}$. Equation (2.2.21) can be realised as a stochastic evolution equation on the triple $H_0^1 \hookrightarrow L_2 \hookrightarrow H^{-1}$. One can easily check that under Assumptions 2.1.1, 2.1.2, and 2.2.1, Assumption 1.0.1 is satisfied, and therefore equation (2.2.21) has a unique L_2-solution $(u_t^{n,m})_{t\in[0,T]}$. We also have that for $n' \geq n$, $f^{n',m} \geq f^{n,m}$. By Theorem 2.2.2 we get that almost surely, for all $t \in [0, T]$

$$u_t^{n',m}(x) \geq u_t^{n,m}(x), \quad \text{for almost every } x. \tag{2.2.22}$$

We define now the stopping time

$$\tau^{R,m} := \inf\{t \geq 0 : \int_Q (u_t^{1,m} + R)_-^2 \, dx > 0\} \wedge T.$$

We claim that for each $R \in \mathbb{N}$, there exists a set Ω_R of full probability, such that for each $\omega \in \Omega_R$, and for all $n \geq R$ we have that

$$u_t^{n,m} = u_t^{R,m}, \quad \text{for } t \in [0, \tau^{R,m}]. \tag{2.2.23}$$

Notice that by (2.2.22) and the definition of $\tau^{R,m}$, for all $n \geq R$

$$f_t^{n,m}(x, u_t^{n,m}(x)) = f_t^{R,m}(x, u_t^{n,m}(x)), \quad \text{for } t \in [0, \tau^{R,m}].$$

This means that for all $n \geq R$ the processes $u_t^{n,m}$ satisfies

$$\begin{aligned} dv_t &= (L_t v_t + f_t^{R,m}(v_t))dt + (M_t^k v_t + g_t^k)dw_t^k, \\ v_0 &= \psi, \end{aligned} \tag{2.2.24}$$

on $[0, \tau^{R,m}]$. The uniqueness of the L_2-solution of the above equation shows (2.2.23). Notice that by Assumption 2.2.1 (iii) and (iv), Theorem 2.1.1 guarantees that $u^{1,m}$ is almost surely essentially bounded in (t, x). Therefore, for almost every $\omega \in \Omega$, $\tau^{R,m} = T$ for all R large enough. On the set $\tilde{\Omega} := \cap_{R \in \mathbb{N}} \Omega_R$ we define $u_t^{\infty,m} = \lim_{n\to\infty} u_t^{n,m}$, where the limit is in the sense of $L_2(Q)$. Since for each $\omega \in \tilde{\Omega}$, we have $u_t^{\infty,m} = u_t^{n,m}$ for all $t \leq \tau^{R,m}$, and for any $n \geq R$, it follows that the process $(u_t^{\infty,m})_{t\in[0,T]}$ is an adapted continuous $L_2(Q)-$valued process such that

i) $u_t^{\infty,m} \in H_0^1$, for $dP \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$

ii) $\int_0^T |u_t^{\infty,m}|_2^2 + |\nabla u_t^{\infty,m}|_2^2 dt < \infty$(a.s.)

iii) $u_t^{\infty,m}$ is almost surely essentially bounded in (t, x)

iv) for all $\phi \in C_c^\infty(Q)$ we have with probability one

$$(u_t^{\infty,m}, \phi) = \int_0^t (a_s^{ij}\partial_{ij}u_s^{\infty,m}, \phi) + (b_s^i \partial_i u_s^{\infty,m} + c_s u_s^m, \phi) + (f_s^m(u_s^{\infty,m}), \phi)ds$$

$$+ \int_0^t (\sigma_s^{ik}\partial_i u_s^{\infty,m} + \nu_s^k u_s^{\infty,m} + g_s^k, \phi)dw_s^k + (\psi, \phi),$$

for all $t \in [0, T]$, where

$$f_t^m(x, r) = \begin{cases} f_t(x, m) & \text{if } r > m \\ f_t(x, r) & \text{if } r \leq m. \end{cases}$$

Now we will let $m \to \infty$. Let us define the stopping time

$$\tau^R := \inf\{t \geq 0 : \int_Q (u_t^{\infty,1} - R)_+^2 dx > 0\} \wedge T.$$

As before we claim that for any $R > 0$, there exists a set Ω'_R of full probability, such that for any $\omega \in \Omega'_R$ and any $m, m' \geq R$,

$$u_t^{\infty,m'} = u_t^{\infty,m} \quad \text{on } [0, \tau^R]. \tag{2.2.25}$$

To show this it suffices to show that for each $R \in \mathbb{N}$, almost surely, for all $m \geq R$, we have $u_t^{n,m} = u_t^{n,R}$ on $[0, \tau^R]$ for all $n \in \mathbb{N}$. To show this we set

$$\tau_n^R := \inf\{t \geq 0 : \int_Q (u_t^{n,1} - R)_+^2 dx > 0\} \wedge T.$$

For all $m \geq R$ we have that the processes $u_t^{n,m}$ satisfy the equation

$$\begin{aligned} dv_t &= (L_t v_t + f_t^{n,R}(v_t))dt + (M_t^k v_t + g_t^k\} dw_t^k, \\ v_0(x) &= \psi(x), \end{aligned} \tag{2.2.26}$$

for $t \leq \tau_n^R$. It follows that almost surely, $u_t^{n,m} = u_t^{n,R}$ for $t \leq \tau_n^R$, for all n. We just note here that by the comparison principle again, we have $\tau^R \leq \tau_n^R$ and this shows (2.2.25). Also for almost every $\omega \in \Omega$, we have $\tau^R = T$ for R large enough. Hence we can define $u_t = \lim_{m \to \infty} u_t^{\infty,m}$, and then one can easily see that u_t has the desired properties.

For the uniqueness, let $u^{(1)}$ and $u^{(2)}$ be solutions of (2.2.19). Then one can define the stopping time

$$\tau_N = \inf\{t \geq 0 : \int_Q (|u_t^{(1)}| - N)_+^2 dx \vee \int_Q (|u_t^{(2)}| - N)_+^2 dx > 0\},$$

to see that for $t \leq \tau_N$, the two solutions satisfy equation (2.2.21) with $n = m = N$, and the claim follows, since $\tau_N = T$ almost surely, for large enough N.

□

2.3 Local supremum estimates

Contrary to [4] where the De Giorgi iteration was used and adapted to the stochastic setting, here, like in Section 2.1, we will use Moser's iteration. This approach has the advantage of providing moment estimates but the proof is somewhat technically more difficult and requires an additional technical assumption, see Assumption 2.3.2 below.

For the sake of clarity we now include only the leading order terms in both the drift and the diffusion, that is, we consider

$$du_t = L_t u_t \, dt + M_t^k u_t \, dw_t^k, \tag{2.3.27}$$

with

$$L_t \varphi = \partial_i (a_t^{ij} \partial_j \varphi), \quad M_t^k \varphi = \sigma_t^{ik} \partial_i \varphi,$$

on G, with the notations $G_R = [4 - R^2, 4] \times B_R$, and $G = G_2$. Since in the following we deal with local properties, restricting our attention to G is not a loss of generality. We will also use the notation $\gamma = (d + 2)/d$ and note that in Lemma 1.2.1 one can choose $r = q = 2\gamma$.

Assumption 2.3.1. For $i, j \in \{1, ..., d\}$, the functions $a^{ij} = a_t^{ij}(x)(\omega)$ and $\sigma^i = (\sigma_t^{ik}(x)(\omega))_{k=1}^{\infty}$ are $\mathcal{P} \times \mathcal{B}(B_2)$-measurable functions on $\Omega \times [0, \infty) \times B_2$ with values in $\mathbb{R}$ and l_2, respectively, bounded by a constant K, such that

$$(2a^{ij} - \sigma^{ik}\sigma^{jk})z_i z_j \geq \lambda |z|^2$$

for a $\lambda > 0$ and for any $z = (z_1, \ldots, z_d) \in \mathbb{R}^d$.

Assumption 2.3.2. For all $p > 1$, $q > 1$,

$$E\|u\|_{p,G}^q < \infty.$$

Notice that, due to the lack of initial or boundary condition, we are not formally in the framework of Theorem 1.0.1. Nevertheless, the concept of solution is defined analogously:

Definition 2.3.1. We will say that $(u_t)_{t\in[0,4]}$ satisfies (or is a solution of) (2.3.27), if u is a strongly continuous $L_2(B_2)$-valued process, $u_t \in H^1(B_2)$ for almost every (t, ω),

$$E \sup_{t\leq 4} |u_t|_2^2 + E \int_0^4 \int_{B_2} |\nabla u_t|^2 dx dt < \infty.$$

and for each $\phi \in C_c^\infty(B_2)$, with probability one,

$$(u_t, \phi) = (u_0, \phi) - \int_0^t (a_t^{ij} \partial_i u_t, \partial_j \phi) dt + \int_0^t (\sigma_t^{ik} \partial_i u_t, \phi) dw_t^k,$$

for all $t \in [0, 4]$.

We start by a weaker supremum estimate, where the uniform norm is estimated in terms of a high L_q-norm, by a localized version of the argument in the previous section.

Lemma 2.3.1. *Let Assumptions 2.3.1-2.3.2 hold. Let $\tau \leq 4$ be a stopping time, u be a solution of (2.3.27) up to τ, and let $f \in C_b^2(\mathbb{R})$, with $ff'' \geq 0$, having bounded first derivative. Then for any $0 < \delta < R \leq 2$ we have*

$$E\|\mathbf{1}_{[0,\tau]} f(u)\|_{\infty, G_{R-\delta}}^q \leq \delta^{2\gamma/(1-\gamma)} C E \|\mathbf{1}_{[0,\tau]} f(u)\|_{q, G_R}^q,$$

where C depends only on d, λ, K.

Proof. Denote $Q = B_R$, $r = 4 - R^2$, let $\varphi \in C_c^\infty(B_R)$, and let $\psi \in C^\infty([4-R^2, 4])$ be an increasing function such that $\psi_r = 0$. Let $\tau' \geq r$ be a stopping time and $\hat{\tau} = \tau \wedge \tau'$. Let us apply Itô's formula to $\int_Q \varphi^2 \psi_t^2 |f(u_t)|^p$. Note that its validity needs to be justified, which can be done by following step-by-step the proof of Lemma 2.1.2 and making use of Assumption 2.3.2 at the passage to the limit. We get

$$
\begin{aligned}
\int_Q \varphi^2 \psi_t^2 |f(u_t)|^p = m_t &+ 2\int_r^t \int_Q \varphi^2 \psi_s \psi_s' |f(u_s)|^p dxds \\
&- \int_r^t \int_Q 2\varphi \partial_i \varphi \psi_s^2 p f(u_s) |f(u_s)|^{p-2} f'(u_s) a_s^{ij} \partial_j u_s dxds \\
&- \int_r^t \int_Q \phi^2 \psi_s^2 p(p-1) |f(u_s)|^{p-2} |f'(u_s)|^2 a_s^{ij} \partial_j u_s \partial_i u_s dxds \\
&- \int_r^t \int_Q \phi^2 \psi_s^2 p f(u_s) |f(u_s)|^{p-2} f''(u_s) a_s^{ij} \partial_j u_s \partial_i u_s dxds \\
&+ \frac{1}{2} \int_r^t \int_Q \varphi^2 \psi_s^2 p(p-1) |f(u_s)|^{p-2} |f'(u_s)|^2 |\sigma^i D_i u_s|^2 dxds \\
&+ \frac{1}{2} \int_r^t \int_Q \varphi^2 \psi_s^2 p f(u_s) |f(u_s)|^{p-2} f''(u_s) |\sigma^i D_i u_s|^2 dxds
\end{aligned}
$$

for $r \leq t \leq \hat{\tau}$, where

$$
m_t = \int_r^t \int_Q \mathbf{1}_{\tau \geq r} \varphi^2 \psi_s^2 p f(u_s) |f(u_s)|^{p-2} f'(u_s) \sigma_s^i \partial_i u_s dxdw_s.
$$

Then by Young's inequality, the parabolicity condition, and the fact that $ff'' \geq 0$, we obtain

$$
\begin{aligned}
\int_Q \varphi^2 \psi_t^2 |f(u_t)|^p \leq m_t &+ 2\int_r^t \int_Q (\varphi^2 \psi_s \psi_s' + |\partial_i \varphi|^2 \psi_s^2) |f(u_s)|^p dxds \\
&- \frac{\lambda}{8} \int_r^t \int_Q \varphi^2 \psi_s^2 p(p-1) |f(u_s)|^{p-2} |f'(u_s)|^2 |\nabla u_s|^2 dxds \quad (2.3.28)
\end{aligned}
$$

for $r \leq t \leq \hat{\tau}$. In particular, (2.3.28) implies that

$$
\begin{aligned}
&E\mathbf{1}_{\tau \geq r} \int_r^{\hat{\tau}} \int_Q p^2 \varphi^2 \psi_s^2 |f(u_s)|^{p-2} |f'(u_s)|^2 |\nabla u_s|^2 dxds \\
&\leq CE\mathbf{1}_{\tau \geq r} \int_r^{\hat{\tau}} \int_Q (\varphi^2 \psi_s \psi_s' + |\partial_i \varphi|^2 \psi_s^2) |f(u_s)|^p dxds,
\end{aligned}
$$

where C depends only on λ. By the Burkholder-Gundy-Davis inequality we have

$$
\begin{aligned}
E\mathbf{1}_{\tau\geq r}\sup_{t\in[r,\hat{\tau}]}|m_t| &\leq CE\mathbf{1}_{\tau\geq r}\left(\int_r^{\hat{\tau}}\left(\int_Q \varphi^2\psi_s^2 pf(u_s)|f(u_s)|^{p-2}f'(u_s)\sigma_s^i\partial_i u_s dx\right)^2 ds\right)^{1/2} \\
&\leq \frac{1}{2}E\mathbf{1}_{\tau\geq r}\sup_{t\in[r,\hat{\tau}]}\int_Q \varphi^2\psi_t^2|f(u_t)|^p dx \\
&\quad + CE\mathbf{1}_{\tau\geq r}\int_r^{\hat{\tau}}\int_Q p^2\varphi^2\psi_s^2|f(u_s)|^{p-2}|f'(u_s)|^2|\nabla u_s|^2 dxds.
\end{aligned}
$$

Hence, by taking suprema in (2.3.28) and using the two estimates above, one can see that

$$
E\mathbf{1}_{\tau\geq r}\sup_{t\in[r,\hat{\tau}]}\int_Q \varphi^2\psi_t^2|f(u_t)|^p dx + E\mathbf{1}_{\tau\geq r}\int_r^{\hat{\tau}}\int_Q p^2\varphi^2\psi_s^2|f(u_s)|^{p-2}|f'(u_s)|^2|\nabla u_s|^2 dxds
$$

$$
\leq CE\mathbf{1}_{\tau\geq r}\int_r^{\hat{\tau}}\int_Q (\varphi^2\psi_s\psi_s' + |\partial_i\varphi|^2\psi_s^2)|f(u_s)|^p dxds,
$$

hence,

$$
E\mathbf{1}_{\tau\geq r}\left[\sup_{t\in[r,v\wedge\tau]}\int_Q \varphi^2\psi_t^2|f(u_t)|^p dx + \int_r^{v\wedge\tau}\int_Q |\nabla(\varphi\psi_s|f(u_s)|^{p/2})|^2 dxds\right]\Bigg|_{v=\tau'}
$$

$$
\leq CE\mathbf{1}_{\tau\geq r}\int_r^{v\wedge\tau}\int_Q (\varphi^2\psi_s\psi_s' + |\partial_i\varphi|^2\psi_s^2)|f(u_s)|^p dxds\Bigg|_{v=\tau'}.
$$

Lemma 1.2.3 therefore can be applied, and we obtain for $q > p$

$$
E\mathbf{1}_{\tau\geq r}\left(\sup_{t\in[r,\tau]}\int_Q \varphi^2\psi_t^2|f(u_t)|^p dx + \int_r^{\tau}\int_Q |\nabla(\varphi\psi_s|f(u_s)|^{p/2})|^2 dxds\right)^{q/p}
$$

$$
\leq C^{q/p}(p/q)^{q/p}\frac{p}{p-q}E\mathbf{1}_{\tau\geq r}\left(\int_r^{\tau}\int_Q (\varphi^2\psi_s\psi_s' + |\partial_i\varphi|^2\psi_s^2)|f(u_s)|^p dxds\right)^{q/p}.
$$

By Lemma 1.2.1 we can estimate the left-hand side from below, and we get

$$
E\mathbf{1}_{\tau\geq r}\left(\int_r^{\tau}\int_Q \varphi^{2\gamma}\psi^{2\gamma}|f(u_t)|^{p\gamma} dxdt\right)^{q/\gamma p}
$$

$$
\leq C^{q/p}(p/q)^{q/p}\frac{p}{p-q}E\mathbf{1}_{\tau\geq r}\left(\int_r^{\tau}\int_Q (\varphi^2\psi_s\psi_s' + |\partial_i\varphi|^2\psi_s^2)|f(u_s)|^p dxds\right)^{q/p}.
$$

We take $\varphi = \varphi_n$, with $|\nabla\varphi_n| \leq C\delta^{-1}2^n$, such that $\varphi_n = 1$ on $B_{R-\delta+2^{-(n+1)}\delta}$ and $\varphi_n = 0$ outside of $B_{R-\delta+2^{-n}\delta}$. Similarly, we take $\psi = \psi_n$ with $|\nabla\psi| \leq$

$C\delta^{-2}2^{2n}$ such that $\psi = 1$ on $[t_0 - (R-\delta)^2 - 2^{-2(n+1)}\delta^2, t_0]$ and $\psi = 0$ outside of $[t_0 - (R-\delta)^2 - 2^{-2n}\delta^2, t_0]$. Let us also introduce the notation $F_n = [t_0 - (R - \delta)^2 - 2^{-2n}\delta^2, t_0] \times B_{R-\delta+2^{-n}\delta}$. Then if we apply the above estimate with $p_n = q\gamma^n$ we have,

$$E\|\mathbf{1}_{\tau\geq r}\mathbf{1}_{[0,\tau]}f(u)\|^q_{p_{n+1},F_{n+1}} \leq (\frac{1}{\delta^2}4^n C\gamma^n)^{1/\gamma^n}\frac{\gamma^n}{\gamma^n-1}E\|\mathbf{1}_{\tau\geq r}\mathbf{1}_{[0,\tau]}f(u)\|^q_{p_n,F_n}$$

By iteration, noting that $F_0 \subset G_R$ and $G_{R-\delta} \subset \cap F_n$, we get the desired estimate, for $\mathbf{1}_{\tau\geq r}f(u)$ instead of $f(u)$. Notice that the fact that the product of the prefactors on the right-hand side, for $n = 1, \ldots$, is finite, is justified in Section 2.1. Finally, notice that

$$E\|\mathbf{1}_{\tau<r}\mathbf{1}_{[0,\tau]}f(u)\|^q_{\infty,G_{R-\delta}} = 0,$$

which finishes the proof. □

Now the main local supremum estimate reads as follows.

Theorem 2.3.2. *Let the conditions of Lemma 2.3.1 be satisfied. Then*

$$E\|\mathbf{1}_{[0,\tau]}f(u)\|^q_{\infty,G_1} \leq q^{aq/2}CE\|\mathbf{1}_{[0,\tau]}f(u)\|^q_{2,G_{3/2}},$$

for constants $a, C > 0$ depending only on d, λ, K.

Proof. Let us denote

$$A(R) = E\|\mathbf{1}_{[0,\tau]}f(u)\|^q_{\infty,G_R}, \quad B(R) = E\|\mathbf{1}_{[0,\tau]}f(u)\|^q_{2,G_R}.$$

By Lemma 2.3.1 and Hölder's inequality we have, with the notation $a' = 2\gamma/(1-\gamma)$

$$A(R) \leq \delta^{a'}CA(R+\delta)^{(q-2)/q}B(R+\delta)^{2/q} \leq \delta^{a'}CA(R+\delta)^{(q-2)/q}B(3/2)^{2/q}, \quad (2.3.29)$$

whenever $R+\delta \leq 3/2$. Now let us choose $\delta = \delta_n$ and $R = R_n = 1 + \sum_{i=1}^n \delta_i$, for $n = 0, 1, \ldots$, such that $R_n \leq 3/2$. Upon iterating (2.3.29), we find

$$A(1) \leq \prod_{i=1}^n (C\delta_i)^{a'[(q-2)/q]^{i-1}} A(R_n)^{[(q-2)/q]^n} B(3/2)^{(2/q)\sum_{i=0}^{n-1}[(q-2)/q]^i}.$$

The exponent of the second term tends to 0, while the exponent of the third term tends to 1. Since $A(R_n) \leq A(3/2) < \infty$, we obtain,

$$A(1) \leq \prod_{i=1}^\infty (C\delta_i)^{a'[(q-2)/q]^{i-1}} B(3/2).$$

Let us choose $\delta_i = i^{-2}/4$. Then

$$\prod_{i=1}^{\infty}(C\delta_i)^{a'[(q-2)/q]^{i-1}} = \exp\{-2\rho a' \sum_{i=1}^{\infty} \frac{\log i}{\rho^i} - 2\rho a \sum_{i=1}^{\infty} \frac{\log C}{\rho^i}\} \le \exp\{a/2 \sum_{i=1}^{\infty} \frac{\log i}{\rho^i}\}$$

for some $a > 0$, where $\rho = q/(q-2)$ and therefore $1/\rho = 1 - 2/q$. The function $h(t) = \log t/\rho^t$ has a unique maximum on $[1,\infty]$, therefore

$$\begin{aligned}
\sum_{i=1}^{\infty} \frac{\log i}{\rho^i} &\le \int_1^{\infty} \frac{\log t}{\rho^t} dt + \max_{t\in[1,\infty]} h(t) \\
&\le \left[\frac{\log t}{\rho^t \log \frac{1}{\rho}}\right]_{t=1}^{\infty} + \frac{1}{\log\rho}\int_1^{\infty} \frac{1}{t\rho^t} dt + \max_{t\in[1,\infty]} \frac{t}{\rho^t} \\
&\le 0 + \frac{1}{\log\rho}\sum_{i=1}^{\infty} \frac{1}{i\rho^i} + \frac{e}{\log\rho} \\
&= \frac{-\log(1-1/\rho)}{\log\rho} + \frac{e}{\log\rho} \le 2\frac{q}{2}\log\frac{q}{2}.
\end{aligned}$$

Therefore,

$$A(1) \le q^{aq/2} B(3/2),$$

which is what we wanted to prove. □

Corollary 2.3.3. *Let the conditions of Lemma 2.3.1 be satisfied with $\tau \equiv 4$. Then for any $n > 1$ and $\alpha > 0$,*

$$P(\|f(u)\|^2_{\infty,G_1} \ge n\alpha, \|f(u)\|^2_{2,G_{3/2}} \le \alpha) \le Ce^{-n^{1/a}}$$

for constants $a, C > 0$ depending only on d, λ, K.

Proof. By Lemma 2.3.2, the processes

$$X_t = \|\mathbf{1}_{[0,t]} f(u)\|^q_{p,G_1}, \;\; A_t = Cq^{aq/2}\|\mathbf{1}_{[0,t]} f(u)\|^q_{2,G_{3/2}}.$$

satisfy the conditions of Lemma 1.2.3 for any p, where C can be chosen independently of p for $p \ge p_0 = p_0(q)$. By Lemma 1.2.3,

$$P(\|f(u)\|^2_{p,G_1} \ge n\alpha, \|f(u)\|^2_{2,G_{3/2}} \le \alpha)$$

$$= P(\|f(u)\|^q_{p,G_1} \ge n^{q/2}\alpha^{q/2}, Cq^{aq/2}\|f(u)\|^2_{2,G_{R+2}} \le C\alpha^{q/2}q^{aq/2}) \le C\left(\frac{q^{aq}}{n^q}\right)^{1/2}.$$

Choosing $q = (n/e)^{1/a}$ and letting $p \to \infty$ yields the result. □

Finally, let us consider the case when the initial value is 0. Note that in this case in the proof of Lemma 2.3.1 the time-cutoff function ψ can be omitted. Doing so and repeating the same steps afterwards, we get the following.

Corollary 2.3.4. *Let Assumptions 2.3.1-2.3.2 hold. Let u be a solution of* (2.3.27) *on $[s,r] \subset [0,4]$, let $f \in C_b^2(\mathbb{R})$, with $ff'' \geq 0$, having bounded first derivative, and suppose that $f(v)(s,\cdot) \equiv 0$. Then for any $n > 1$ and $\alpha > 0$,*

$$P(\|f(u)\|^2_{\infty,[s,r]\times B_1} \geq n\alpha, \|f(u)\|^2_{2,[s,r]\times B_2} \leq \alpha) \leq Ce^{-n^{1/a}}$$

for constants $a, C > 0$ depending only on d, λ, K.

2.4 A Harnack inequality and continuity of solutions

Denote by Λ the set of functions v on $[0,4] \times B_2$ such that $v \geq 0$ and

$$|\{x \in B_2 | \, v_0(x) \geq 1\}| \geq \frac{1}{2}|B_2|.$$

Let us recall the Harnack inequality essentially proved in [17]: *If u is a solution of $du = \partial_i(a^{ij}\partial_j u)dt$ and $u \in \Lambda$, then*

$$\inf_{G_1} u \geq h$$

with $h = h(d, \lambda, K) > 0$. In the stochastic case clearly it is not expected that such a lower estimate holds uniformly in ω. It does hold, however, with h above replaced with a strictly positive random variable, this is the assertion of our main theorem.

Theorem 2.4.1. *Let Assumptions 2.3.1-2.3.2 hold. Let u be a solution of* (2.3.27) *such that on an event $A \in \mathscr{F}$, $u \in \Lambda$. Then for any $N > 0$ there exists a set $D \in \mathscr{F}$, with $P(D) \leq Ce^{-N^{1/a}}$, such that on $A \cap D^c$,*

$$\inf_{(t,x)\in G_1} u_t(x) \geq e^{-N}.$$

where C and a, depend only on d, λ and K.

Later on we will refer to the quantity e^{-N} above as the lower bound corresponding to the probability $Ce^{-N^{1/a}}$. We begin with a simple lemma.

Lemma 2.4.2. *For any $c > 0$, there exists $N_0(c) > 0$, such that for any continuous local martingale m_t, and for any $N \geq N_0$,*

$$P\left(\sup_{t\geq 0}(m_t - c\langle m\rangle_t) > N\right) \leq Ce^{-Nc/4},$$

with an absolute constant C.

Proof. Let B be a Wiener process for which $B_{\langle m\rangle_t} = m_t$. Then for any $\beta > 0$

$$P\left(\sup_{t\geq 0}(m_t - c\langle m\rangle_t) > N\right) \leq P\left(\sup_{s\geq 0}(B_s - cs) > N\right)$$

$$\leq P(\sup_{s\in[0,\beta]} B_s > N) + \sum_{i=1}^{\infty} P(\sup_{s\in[0,(i+1)\beta]} B_s > ic\beta).$$

Recall that for any $\alpha > 0$

$$P(\sup_{s\in[0,\beta]} B_s \geq \alpha) \leq \sqrt{\frac{2}{\pi\beta}}\frac{\beta}{\alpha}\int_{\alpha}^{\infty}\frac{x}{\beta}e^{-x^2/2\beta}\,dx = C\sqrt{\frac{\beta}{\alpha^2}}e^{-\alpha^2/2\beta}.$$

Therefore,

$$P\left(\sup_{t\geq 0}(m_t - c\langle m\rangle_t) > N\right) \leq C\sqrt{\frac{\beta}{N^2}}e^{-N^2/2\beta} + \sum_{i=1}^{\infty} C\sqrt{\frac{(i+1)}{c^2i^2\beta}}e^{-c^2i^2\beta/2(i+1)}.$$

Choosing $\beta = N/c$ yields the claim. $\square$

Next, we establish what can be considered a weak version of Theorem 2.4.1.

Lemma 2.4.3. *Let Assumptions 2.3.1-2.3.2 hold. Let u be a solution of* (2.3.27), *such that on $A \in \mathscr{F}$, $u \in \Lambda$. Then for any $N > 0$, there exists a set $D_1 \in \mathscr{F}$, with $P(D_1) \leq Ce^{-cN}$, such that on $A \cap D_1^c$, for all $t \in [0, 4]$,*

$$|\{(x \in B_\rho|\, v(t,x) \geq e^{-N}\}| \geq \frac{1}{8}|B_\rho|,$$

where ρ is defined by

$$|B_\rho| = \frac{3}{4}|B_2|,$$

and the constants $c, C > 0$, depend only on d, λ, K.

Proof. Clearly it is sufficient to prove the statement for $N > N_0$ for some N_0.

Introduce the functions

$$f_h(x) = \begin{cases} a_h x + b_h & \text{if } x < -h/2 \\ \log^+ \frac{1}{x+h} & \text{if } x \geq -h/2, \end{cases}$$

for $h > 0$ where a_h and b_h is chosen such that f_h and f_h' are continuous. Let κ be nonnegative a C^∞ function on $\mathbb{R}$, bounded by 1, supported on $\{|x| < 1\}$, and having unit integral. Denote $\kappa_h(x) = h^{-1}\kappa(x/h)$ and

$$F_h = f_h * \kappa_{h/4}.$$

We claim that F_h has the following properties:

(i) $F_h(x) = 0$ for $x \geq 1$;

(ii) $F_h(x) \leq \log(2/h)$ for $x \geq 0$;

(iii) $F_h(x) \geq \log(1/2h)$ for $x \leq h/2$;

(iv) $F_h \in \mathcal{D}$ and $F_h''(x) \geq (F_h'(x))^2$ for $x \geq 0$.

The first three properties are obvious, while for the last one notice that F_h has bounded second derivative, $f_h''(x) \geq (f_h'(x))^2$ for $x \geq -h/2$, and therefore, for $x \geq 0$

$$\begin{aligned}(F_h'(x))^2 &= \left(\int f_h'(x-z)\kappa_{h/4}^{1/2}(z)\kappa_{h/4}^{1/2}(z)\,dz \right)^2 \\ &\leq \int (f_h'(x-z))^2 \kappa_{h/4}(z)\,dz \\ &\leq \int f_h''(x-z)\kappa_{h/4}(z)\,dz = F_h''(x).\end{aligned}$$

Let us denote $v = F_h(u)$. Applying Itô's formula and using the parabolicity condition, we get

$$\begin{aligned}\int_{B_2} \varphi^2 v_t\,dx - \int_{B_2} \varphi^2 v_0\,dx \leq \int_0^t \int_{B_2} C\varphi\nabla\varphi\nabla v - (\lambda/2)\varphi^2 F_h''(u)(\nabla u)^2\,dx\,ds \\ + \int_0^t \int_{B_2} \varphi^2 M^k v\,dx\,dw_s^k \end{aligned} \tag{2.4.30}$$

for any $\varphi \in C_c^\infty$. Let us denote the stochastic integral above by m_t, and notice that provided $|\varphi| \leq 1$,

$$\langle m \rangle_t \leq C \int_0^t \int_{B_2} \varphi^2 (\nabla v)^2\,dx\,ds.$$

Let c be such that $cC \leq \lambda/4$. From Lemma 2.4.2, there exists a set D_1 with $P(D_1) \leq Ce^{-Nc/4}$, such that on D_1^c we have

$$\int_{B_2} \varphi^2 v_t \, dx - \int_{B_2} \varphi^2 v_0 \, dx$$

$$\leq N + \int_0^t \int_{B_2} C\varphi\nabla\varphi\nabla v - (\lambda/2)\varphi^2 F_h''(u)(\nabla u)^2 + cC\varphi^2(\nabla v)^2 \, dx \, ds. \quad (2.4.31)$$

On $A \cap D_1^c$, by the property (iv) above, we have $F_h''(u)(\nabla u)^2 \geq (\nabla v)^2$, and therefore

$$\int_{B_2} \varphi^2 v_t \, dx \leq N + C\int_{B_2} |\nabla\varphi|^2 \, dx + \int_{B_2} \varphi^2 v_0 \, dx. \quad (2.4.32)$$

Let us denote

$$\mathcal{O}_t(h) = \{x \in B_\rho : u(t,x) \geq h\}.$$

Choosing φ to be 1 on B_ρ, by properties (i), (ii), and (iii) of F_h and (2.4.32), on $A \cap D_1^c$, for all $t \in [0,4]$

$$|B_\rho \setminus \mathcal{O}_t(h/2)| \log(1/2h) \leq C + N + \frac{1}{2}\log(2/h)|B_2| = C + N + \frac{2}{3}\log(2/h)|B_\rho|.$$

Hence

$$|\mathcal{O}_t(h/2)| \geq |B_\rho| - \frac{C+N}{\log(1/2h)} - \frac{2}{3}\frac{\log(2/h)}{\log(1/2h)}|B_\rho|,$$

and choosing $N_0 = C$ and $h = 2e^{-C'N}$ for a sufficiently large C' finishes the proof of the lemma. □

Proof of Theorem 2.4.1

By Lemma 2.4.3, there exists a set D_1 with $P(D_1) \leq Ce^{-cN}$ such that on $A \cap D_1^c$ we have

$$|\{(x \in B_\rho | \, v(t,x) \geq e^{-N}| \geq \frac{1}{8}|B_\rho|, \quad (2.4.33)$$

for all $t \in [0,4]$. Let us denote $h := e^{-N}$. For $0 < \epsilon \leq h/2$, we introduce the function

$$f_\epsilon(x) = \begin{cases} a_\epsilon x + b_\epsilon & \text{if } x < -\epsilon/2 \\ \log^+ \frac{h}{x+\epsilon} & \text{if } x \geq -\epsilon/2, \end{cases}$$

where a_ϵ and b_ϵ is chosen such that f_ϵ and f_ϵ' are continuous. Let κ be a nonnegative C^∞ function on $\mathbb{R}$, bounded by 1, supported on $\{|x| < 1\}$, and having unit integral. Denote $\kappa_\varepsilon(x) = \varepsilon^{-1}\kappa(x/\varepsilon)$ and

$$F_\epsilon = f_\epsilon * \kappa_{\epsilon/4}.$$

Similarly to F_h in the proof of Lemma 2.4.3, F_ϵ has the following properties:

(i) $F_\epsilon(x) = 0$ for $x \geq h$;

(ii) $F_\epsilon(x) \leq \log(2h/\epsilon)$ for $x \geq 0$;

(iii) $F_\epsilon(x) \geq \log(h/(x+\epsilon)) - 1$ for $x \geq 0$;

(iv) $F_\epsilon \in \mathcal{D}$ and $F''_\epsilon(x) \geq (F'_\epsilon(x))^2$ for $x \geq 0$.

Let us denote $v = F_\epsilon(u)$. Similarly to (2.4.31), there exists a set D_2 with $P(D_2) \leq Ce^{-Nc}$, such that on D_2^c we have

$$\int_{B_2} \varphi^2 v_t \, dx - \int_{B_2} \varphi^2 v_0 \, dx$$
$$\leq N + \int_0^t \int_{B_2} C\varphi \nabla\varphi \nabla v - (\lambda/2)\varphi^2 F''_\epsilon(u)(\nabla u)^2 + (\lambda/4)\varphi^2(\nabla v)^2 \, dx \, ds.$$

On $A \cap D_2^c$, by property (iv), we have,

$$\int_0^4 \int_{B_2} \varphi^2 |\nabla v_t|^2 \, dxdt \leq C(N + \int_{B_2} |\nabla\varphi|^2 \, dx + \int_{B_2} \varphi^2 v_2 \, dx). \qquad (2.4.34)$$

By choosing $\varphi \in C_c^\infty(B_2)$ with $0 \leq \varphi \leq 1$ and $\varphi = 1$ on B_ρ we get,

$$\int_0^4 \int_{B_\rho} |\nabla v_t|^2 \, dxdt \leq C(N + \int_{B_2} |\nabla\varphi|^2 \, dx + \int_{B_2} \varphi^2 v_0 \, dx).$$

Hence, by property (ii),

$$\int_0^4 \int_{B_\rho} |\nabla v_t|^2 \, dxdt \leq CN + C + C\log\frac{2h}{\epsilon}. \qquad (2.4.35)$$

Using property (i), by Lemma 1.2.2 we get for all t

$$\int_{B_\rho} |v_t|^2 dx \leq C\frac{\rho^{2(d+1)}}{|O_t(h)|^2} \int_{B_\rho} |\nabla v_t|^2 dx,$$

which, by virtue of (2.4.33) and (2.4.35) implies

$$\int_0^4 \int_{B_\rho} |v_t|^2 dx \leq C + CN + C\log\frac{2h}{\epsilon}.$$

on $A \cap D_1^c \cap D_2^c$. By Corollary 2.3.3 and noting that $G_{3/2} \subset [0,4] \times B_\rho$ we get that there exists a set $D_3 \in \mathscr{F}$ with $P(D_3) \leq Ce^{-N^{1/a}}$, such that on $A \cap D_1^c \cap D_2^c \cap D_3^c$

we have

$$\sup_{(t,x)\in G_1} v_t(x) \leq [N(C+CN+C\log\frac{2h}{\epsilon})]^{1/2}.$$

By applying property (iii), we get

$$\sup_{(t,x)\in G_1} \log\frac{h}{u_t(x)+\epsilon} \leq [N(C+CN+C\log\frac{2h}{\epsilon})]^{1/2}+1,$$

and therefore,

$$\inf_{(t,x)\in G_1} u_t(x) \geq he^{-[N(C+CN+C\log 2h - C\log\epsilon)]^{1/2}-1} - \epsilon.$$

Letting $\epsilon = e^{-c'N}$ with a sufficiently large c', it is easy to see that the right-hand side above is bounded from below by ϵ, finishing the proof. $\square$

Finally let us present an application of Theorem 2.4.1 which asserts the pointwise continuity of solutions. In particular, we find that the set of discontinuity points of the solution is a.s. of first category and has measure 0.

Theorem 2.4.4. *Let Assumptions 2.3.1-2.3.2 hold. Let u be a solution of* (2.3.27) *and* $(t_0, x_0) \in (0,4)\times B_2$. *Then u is almost surely continuous at* (t_0, x_0).

Proof. Consider the parabolic transformations $\mathfrak{P}_{\alpha,t',x'}$:

$$t \to \alpha^2 t + t',$$

$$x \to \alpha x + x'.$$

It is easy to see that if v is a solution of (2.3.27) on a cylinder Q, then $v\circ\mathfrak{P}^{-1}_{\alpha,t',x'}$ is also solution of (2.3.27), on the cylinder $\mathfrak{P}_{\alpha,t',x'}Q$, with another sequence of Wiener martingales on another filtration, and with different coefficients that still satisfy Assumption 2.3.1 with the same bounds. To ease notation, for a cylinder Q let $\mathfrak{P}_Q$ denote the unique parabolic transformation that maps Q to G, if such exists. Also, for an interval $[s,r]\subset[0,4]$ let $\mathfrak{P}_{[s,r]} = \mathfrak{P}_{2/\sqrt{r-s},-4s/(r-s),0}$. That is, $\mathfrak{P}_{[s,r]}[s,r]\times B_1 = [0,4]\times B_{2/\sqrt{r-s}}$, which, when $r-s\leq 1$, contains G.

Without loss of generality $x_0 = 0$ can and will be assumed, as will the almost sure boundedness of u on G, since these can be achieved with appropriate parabolic transformations, using the boundedness obtained on sub-cylinders in Theorem 2.3.2. Also let us fix a probability $\delta > 0$, denote the corresponding lower bound $3\epsilon_2$ obtained from the Harnack inequality, and take an arbitrary $0 < \epsilon_1 < \epsilon_2/2$.

Let us consider a smooth version of the function $(\cdot)^+$. That is, take a convex

$f \in C^\infty$ such that

(i) $f(t) = 0$ for $t \leq 0$;

(ii) $f(t) \leq t$;

(iii) $f(t) \leq \epsilon_1^2/6$ only if $t \leq \epsilon_1^2/6$.

Apply Theorem 2.3.4 twice with the function f, with the interval $[t_0 - 4s, t_0 + s]$, and with solutions $v = u - \sup_{\{t_0-4s\}\times B_2} u$ and $v = -u + \inf_{\{t_0-4s\}\times B_2} u$. Also notice that (for both choices of v)

$$\|f(v)\|^2_{2,[t_0-4s,t_0+s]\times B_2} \leq Cs\|u\|^2_{\infty,G} \to 0$$

as $s \to 0$ for almost every ω, and thus in probability as well, in other words,

$$P(\|f(v)\|^2_{2,[t_0-4s,t_0+s]\times B_2} > \alpha)$$

can be made arbitrarily small by choosing s sufficiently small. Therefore, we obtain an $s > 0$ and an event Ω_0, with $P(\Omega_0) > 1 - \delta$, such that on Ω_0,

$$\sup_{[t_0-4s,t_0+s]\times B_1} u - \sup_{\{t_0-4s\}\times B_2} u < \epsilon_1^2/6$$

$$\inf_{[t_0-4s,t_0+s]\times B_1} u - \inf_{\{t_0-4s\}\times B_2} u > -\epsilon_1^2/6.$$

Let us rescale u at the starting time:

$$u'_\pm(t,x) = \pm\left(2\frac{u(t,x) - \sup_{\{t_0-4s\}\times B_2} u}{\sup_{\{t_0-4s\}\times B_2} u - \inf_{\{t_0-4s\}\times B_2} u} + 1\right),$$

that is, $\sup_{B_2} u'_\pm(t_0 - 4s, \cdot) = 1, \inf_{B_2} u'_\pm(t_0 - 4s, \cdot) = -1$. Now we can write $\Omega_0 = \Omega_A \cup \Omega_B$, where

- On Ω_A, $\operatorname{osc}_{\{t_0-4s\}\times B_2} u < \epsilon_1/3$, and therefore, $\operatorname{osc}_{[t_0-4s,t_0+s]\times B_1} u < \epsilon_1/3 + 2\epsilon_1^2/6 < \epsilon_1$;

- On Ω_B, $|u'_\pm| < 1 + 2(\epsilon_1^2/6)/(\epsilon_1/3) = 1 + \epsilon_1$, on $[t_0 - 4s, t_0 + s] \times B_1$.

Notice that in the event Ω_B, on the cylinder $[t_0 - 4s, t_0 + s] \times B_1$, the functions $u'_\pm/(1+\epsilon_1)+1$ take values between 0 and 2. Therefore one of $(u'_\pm/(1+\epsilon_1)+1) \circ \mathfrak{P}^{-1}_{[t_0-4s,t_0+s]}\Big|_G$, denoted for the moment by u'', satisfies the conditions of Theorem 2.4.1 with $A = \Omega_B$.

We obtain that on an event Ω'_B

$$\inf_{G_1} u'' > 3\epsilon_2,$$

and thus

$$\mathrm{osc}_Q u < \frac{(2-3\epsilon_2)(1+\epsilon_1)}{2}\mathrm{osc}_{\{t_0-4s\}\times B_2} u < (1-\epsilon_2)\mathrm{osc}_{\{t_0-4s\}\times B_2} u,$$

where $Q = \mathfrak{P}^{-1}_{[t_0-4s,t_0+s]} G_1$. Moreover, $P(\Omega_B \setminus \Omega'_B) < \delta$. Also, notice that $(t_0, 0) \in Q$. Let us denote $\Omega_1 = \Omega_A \cup \Omega'_B$. We have shown the following lemma:

Lemma 2.4.5. *Let $\delta > 0$ and let $3\epsilon_2$ be the lower bound corresponding to the probability δ obtained from the Harnack inequality. For any u that is a solution of (2.3.27) on G, $t_0 > 0$, and for any sufficiently small $\epsilon_1 > 0$ there exists an $s > 0$ and an event Ω_1 such that*

(i) $P(\Omega_1) > 1 - 2\delta$;

(ii) On Ω_1, at least one of the following is satisfied:

- *(a)* $\mathit{osc}_Q u < \epsilon_1$;
- *(b)* $\mathit{osc}_Q u < (1-\epsilon_2)\mathit{osc}_G u$,

where $Q = \mathfrak{P}^{-1}_{[t_0-4s,t_0+s]}(G_1)$.

Now take $u = u^{(0)}$ and $t_0 = t_0^{(0)}$ from the statement of the theorem and a sequence $(\epsilon_1^{(n)})_{n=0}^{\infty} \downarrow 0$, and for $n \geq 0$ proceed inductively as follows:

- Apply Lemma 2.4.5 with $u^{(n)}$, $t_0^{(n)}$, and $\epsilon_1^{(n)}$, and take the resulting $\Omega_1^{(n)}$ and $Q^{(n)}$;
- Let $u^{(n+1)} = u^{(n)} \circ \mathfrak{P}^{-1}_{Q^{(n)}}$ and $(t_0^{(n+1)}, 0) = \mathfrak{P}_{Q^{(n)}}(t_0^{(n)}, 0)$.

On $\limsup_{n\to\infty} \Omega_1^{(n)}$ the function u is continuous at the point $(t_0, 0)$. Indeed, the sequence of cylinders $Q^{(0)}, \mathfrak{P}^{-1}_{Q^{(0)}} Q^{(1)}, \mathfrak{P}^{-1}_{Q^{(0)}} \mathfrak{P}^{-1}_{Q^{(1)}} Q^{(2)}, \ldots$ contain $(t_0, 0)$, and the oscillation of u on these cylinders tends to 0. However, $P(\limsup_{n\to\infty} \Omega_1^{(n)}) \geq 1 - 2\delta$, and since δ can be chosen arbitrarily small, u is continuous at $(t_0, 0)$ with probability 1, and the proof is finished.

□

Chapter 3

Degenerate equations - solvability

The condition $\lambda > 0$ in Assumption 1.0.1 is crucial, for example the smoothing property expressed by Theorem 1.0.1 can clearly not be expected to hold otherwise. However, degenerating operators, i.e. ones for which the coercivity condition holds only with $\lambda = 0$ arise naturally from an important application of SPDEs, the Zakai equation for the nonlinear filtering. Their solvability in W^m_p spaces has been claimed first in [26]. However, the proof, in particular, the a priori estimate for each partial derivative contained a nontrivial gap for the $p \neq 2$ case. It turns out that it is actually not possible to estimate each partial derivative separately, but one has to view the vector of derivatives as a whole, and estimate it using the vector-valued equation it satisfies. This motivates to consider systems of equations in the first place, and leads to some interesting differences from the scalar case. We note that a quite different approach to investigate what the "appropriate" stochastic parabolicity condition is for systems of equations can be found in [30], with the attention restricted to the L_2 scale and constant coefficients. We also note that in the nondegenerate case a complete theory of SPDEs in W^m_p spaces is established in [18]. One rationale behind solving equations in W^m_p for large p is the following. By Sobolev embedding, the solution is n times continuously differentiable if it is in W^m_p with $m - d/p > n$. On the other hand it is expected that solvability in W^m_p requires (roughly) $m = n + d/p + \varepsilon$ bounded derivatives from the coefficients. So in order to relax the regularity assumptions on the coefficients, one wishes to choose p sufficiently large. The content of this chapter is based on the author's joint work with István Gyöngy and Nicolai Krylov, in the paper [11].

3.1 Formulation

Let $M \geq 1$ be an integer, and let $\langle \cdot, \cdot \rangle$ and $\langle \cdot \rangle$ denote the scalar product and the norm in $\mathbb{R}^M$, respectively. By $\mathbb{T}^M$ we denote the set of $M \times M$ matrices, which we consider as a Euclidean space $\mathbb{R}^{M^2}$. For an integer $m \geq 1$ we define $l_2(\mathbb{R}^m)$ as the space of sequences $\nu = (\nu^1, \nu^2, ...)$ with $\nu^k \in \mathbb{R}^m$, $k \geq 1$, and finite norm

$$\|\nu\|_{l_2} = \Big(\sum_{k=1}^{\infty} \langle \nu^k \rangle^2\Big)^{1/2}$$

We look for $\mathbb{R}^M$-valued functions $u_t(x) = (u_t^1(x), ..., u_t^M(x))$, of $\omega \in \Omega$, $t \in [0, T]$ and $x \in \mathbb{R}^d$, which satisfy the system of equations

$$\begin{aligned} du_t =& [a_t^{ij} D_{ij} u_t + b_t^i D_i u_t + c u_t + f_t]\, dt \\ &+ [\sigma_t^{ik} D_i u_t + \nu_t^k u_t + g_t^k]\, dw_t^k, \end{aligned} \tag{3.1.1}$$

and the initial condition

$$u_0 = \psi, \tag{3.1.2}$$

where $a_t = (a_t^{ij}(x))$ takes values in the set of $d \times d$ symmetric matrices,

$$\sigma_t^i = (\sigma_t^{ik}(x))_{k=1}^{\infty} \in l_2, \quad b_t^i(x) \in \mathbb{T}^M, \quad c_t(x) \in \mathbb{T}^M,$$

$$\nu_t(x) \in l_2(\mathbb{T}^M), \quad f_t(x) \in \mathbb{R}^M, \quad g_t(x) \in l_2(\mathbb{R}^M) \tag{3.1.3}$$

for $i = 1, ..., d$, for all $\omega \in \Omega$, $t \geq 0$, $x \in \mathbb{R}^d$. Note that with the exception of a^{ij} and σ^{ik}, all 'coefficients' in equation (3.1.1) mix the coordinates of the process u.

Let m be a nonnegative integer, $p \in [2, \infty)$ and make the following assumptions.

Assumption 3.1.1. The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m, 2)$ and of b^i and c up to order m are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions, in magnitude bounded by K for all $i, j \in \{1, 2, ..., d\}$. The derivatives in x of the l_2-valued functions $\sigma^i = (\sigma^{ik})_{k=1}^{\infty}$ and the $l_2(\mathbb{T}^M)$-valued function ν up to order $m + 1$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued and $l_2(\mathbb{T}^M)$-valued functions, respectively, in magnitude bounded by K.

Assumption 3.1.2. The free data, $(f_t)_{t\in[0,T]}$ and $(g_t)_{t\in[0,T]}$ are predictable processes with values in $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$ and $W_p^{m+1}(\mathbb{R}^d, l_2(\mathbb{R}^M))$, respectively, such that almost surely

$$\mathcal{K}^p_{m,p}(T) = \int_0^T \left(|f_t|^p_{W^m_p} + |g_t|^p_{W^{m+1}_p} \right) dt < \infty. \tag{3.1.4}$$

The initial value, ψ is an $\mathcal{F}_0$-measurable $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued random variable.

To formulate the parabolicity condition for the system, set

$$\alpha^{ij} = 2a^{ij} - \sigma^{ik}\sigma^{jk} \quad i, j = 1, \dots, d$$

and

$$\beta^i = b^i - \sigma^{ir}\nu^r, \quad i = 1, \dots, d.$$

Assumption 3.1.3. There exist a constant $K_0 > 0$ and a $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable $\mathbb{R}^M$-valued bounded function $h = (h^i_t(x))$, whose first order derivatives in x are bounded functions, such that for all $\omega \in \Omega$, $t \geq 0$ and $x \in \mathbb{R}^d$

$$|h| + |Dh| \leq K, \tag{3.1.5}$$

and for all $(\lambda_1, ..., \lambda_d) \in \mathbb{R}^d$

$$|\sum_{i=1}^d (\beta^{ikl} - \delta^{kl} h^i)\lambda_i|^2 \leq K_0 \sum_{i,j=1}^d \alpha^{ij}\lambda_i\lambda_j \quad \text{for } k, l = 1, ..., M. \tag{3.1.6}$$

Remark 3.1.1. Let Assumption 3.1.1 hold with $m = 0$ and the first order derivatives of b^i in x are bounded by K for each $i = 1, 2, ...d$. Then notice that condition (3.1.6) is a natural extension of the degenerate parabolicity condition to systems of stochastic PDEs. Indeed, when $M = 1$ then taking $h^i = \beta^i$ for $i = 1, ..., d$, we can see that Assumption 3.1.3 is equivalent to $\alpha \geq 0$. Let us analyse now Assumption 3.1.3 for arbitrary $M \geq 1$. Notice that it holds when α is uniformly elliptic, i.e., $\alpha \geq \kappa I_d$ with a constant $\kappa > 0$ for all ω, $t \geq 0$ and $x \in \mathbb{R}^d$. Indeed, due to Assumption 3.1.1 there is a constant $N = N(K, d)$ such that

$$|\sum_{i=1}^d (\beta^{ikl} - \delta^{kl} h^i)\lambda_i|^2 \leq N \sum_{i=1}^d |\lambda_i|^2 \quad \text{for every } k, l = 1, 2, ..., M,$$

which together with the uniform ellipticity of α clearly implies (3.1.6). Notice also that (3.1.6) holds in many situations when instead of the strong ellipticity of α we only have $\alpha \geq 0$. Such examples arise, for example, when $a^{ij} = \sigma^{ir}\sigma^{jr}/2$ for all $i, j = 1, ..., d$, and b and ν are such that β^i is a diagonal matrix for each $i = 1, ..., d$, and the diagonal elements together with their first order derivatives in x are bounded by a constant K. As another simple example, consider the system of equations

$$
\begin{aligned}
du_t(x) =& \{\tfrac{1}{2}D^2u_t(x) + Dv_t(x)\}\,dt + \{Du_t(x) + v_t(x)\}\,dw_t \\
dv_t(x) =& \{\tfrac{1}{2}D^2v_t(x) - Du_t(x)\}\,dt + \{Dv_t(x) - u_t(x)\}\,dw_t
\end{aligned}
$$

for $t \in [0, T]$, $x \in \mathbb{R}$, for a 2-dimensional process $(u_t(x), v_t(x))$, where w is a one-dimensional Wiener process. In this example $\alpha = 0$ and $\beta = 0$. Thus clearly, condition (3.1.6) is satisfied.

Later it will be convenient to use condition (3.1.6) in an equivalent form, which we discuss in the next remark.

Remark 3.1.2. Notice that condition (3.1.6) in Assumption 3.1.3 can be reformulated as follows: There exists a constant K_0 such that for all values of the arguments and all continuously differentiable $\mathbb{R}^M$-valued functions $u = u(x)$ on $\mathbb{R}^d$ we have

$$
\langle u, b^i D_i u\rangle - \sigma^{ik}\langle u, \nu^k D_i u\rangle \le K_0 \Big|\sum_{i,j=1}^{d} \alpha^{ij}\langle D_i u, D_j u\rangle\Big|^{1/2}\langle u\rangle + h^i\langle D_i u, u\rangle. \quad (3.1.7)
$$

Indeed, set $\hat{\beta}^i = \beta^i - h^i I_M$, where I_M is the $M \times M$ unit matrix and observe that, (3.1.7) means that

$$
\langle u, \hat{\beta}^i D_i u\rangle \le K_0 \Big|\sum_{i,j=1}^{d} \alpha^{ij}\langle D_i u, D_j u\rangle\Big|^{1/2}\langle u\rangle.
$$

By considering this relation at a fixed point x and noting that then one can choose u and Du independently, we conclude that

$$
\langle \sum_i \hat{\beta}^i D_i u\rangle^2 \le K_0^2 \alpha^{ij}\langle D_i u, D_j u\rangle \quad (3.1.8)
$$

and (3.1.6) follows (with a different K_0) if we take $D_i u^k = \lambda_i \delta^{kl}$.

On the other hand, (3.1.6) means that for any l without summation on l

$$
\Big|\sum_i \hat{\beta}^{ikl} D_i u^l\Big|^2 \le K_0 \alpha^{ij}(D_i u^l) D_j u^l.
$$

But then by Cauchy's inequality similar estimate holds after summation on l is done and carried inside the square on the left-hand side. This yields (3.1.8) (with a different constant K_0) and then leads to (3.1.7).

The notion of solution to (3.1.1)-(3.1.2) is a straightforward adaptation of Definition 1.0.1. Namely, $u = (u^1, ..., u^M)$ is a solution on $[0, \tau]$, for a stopping time $\tau \le T$, if it is a $W^1_p(\mathbb{R}^d, \mathbb{R}^M)$-valued predictable function on $[0, \tau]$,

$$\int_0^\tau |u_t|^p_{W^1_p}\, dt < \infty \quad \text{(a.s.)},$$

and for each $\mathbb{R}^M$-valued $\varphi = (\varphi^1, ..., \varphi^M)$ from $C_0(\mathbb{R}^d)$ with probability one

$$(u_t, \varphi) = (\psi, \varphi) + \int_0^t \{-(a^{ij}_s D_i u_s, D_j\varphi) + (\bar{b}^i_s D_i u_s + c_s u_s + f_s, \varphi)\}\, ds \tag{3.1.9}$$

$$+ \int_0^t (\sigma^{ir}_s D_i u_s + \nu^r_s u_s + g^r(s), \varphi)\, dw^r_s \tag{3.1.10}$$

for all $t \in [0, \tau]$, where $\bar{b}^i = b^i - D_j a^{ij} I_M$. Here, and later on (Ψ, Φ) denotes the inner product in the L_2-space of $\mathbb{R}^M$-valued functions Ψ and Φ defined on $\mathbb{R}^d$. The main result now reads as follows.

Theorem 3.1.1. *Let Assumption 3.1.3 hold. If Assumptions 3.1.1 and 3.1.2 also hold with $m \ge 0$, then there is at most one solution to* (3.1.1)-(3.1.2) *on* $[0, T]$. *If together with Assumption 3.1.3, Assumptions 3.1.1 and 3.1.2 hold with $m \ge 1$, then there is a unique solution $u = (u^l)^M_{l=1}$ to* (3.1.1)-(3.1.2) *on* $[0, T]$. *Moreover, u is a weakly continuous $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process, it is strongly continuous as a $W^{m-1}_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and for every $q > 0$ and $n \in \{0, 1, ..., m\}$*

$$E \sup_{t\in[0,T]} |u_t|^q_{W^n_p} \le N(E|\psi|^q_{W^n_p} + E\mathcal{K}^q_{n,p}(T)) \tag{3.1.11}$$

with $N = N(m, p, q, d, M, K, T)$.

In the case $p = 2$ we present also a modification of Assumption 3.1.3, in order to cover an important class of stochastic PDE systems, the *hyperbolic symmetric systems.*

Observe that if in (3.1.6) we replace β^{ikl} with β^{ilk}, nothing will change. By the convexity of t^2 condition (3.1.6) then holds if we replace β^{ilk} with $(1/2)[\beta^{ilk} + \beta^{ikl}]$. Since

$$|a - b|^2 \le |a + b|^2 + 2a^2 + 2b^2$$

this implies that (3.1.6) also holds for

$$\bar{\beta}^{ikl} = (\beta^{ikl} - \beta^{ilk})/2$$

in place of β^{ikl}, which is the antisymmetric part of $\beta^i = b^i - \sigma^{ir}\nu^r$.

Hence the following condition is weaker than Assumption 3.1.3.

Assumption 3.1.4. There exist a constant $K_0 > 0$ and a $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable $\mathbb{R}^M$-valued function $h = (h_t^i(x))$ such that (3.1.5) holds, and for all $\omega \in \Omega$, $t \geq 0$ and $x \in \mathbb{R}^d$ and for all $(\lambda_1, ..., \lambda_d) \in \mathbb{R}^d$

$$|\sum_{i=1}^{d}(\bar{\beta}^{ikl} - \delta^{kl}h^i)\lambda_i|^2 \leq K_0 \sum_{i,j=1}^{d} \alpha^{ij}\lambda_i\lambda_j \quad \text{for } k, l = 1, ..., M. \tag{3.1.12}$$

The following result in the special case of deterministic PDE systems is indicated and a proof is sketched in [14].

Theorem 3.1.2. *Take $p = 2$ and replace Assumption 3.1.3 with Assumption 3.1.4 in the conditions of Theorem 3.1.1. Then the conclusion of Theorem 3.1.1 holds with $p = 2$.*

Remark 3.1.3. Notice that Assumption 3.1.4 obviously holds with $h^i = 0$ if the matrices β^i are symmetric and $\alpha \geq 0$. When $a = 0$ and $\sigma = 0$ then the system is called a *first order symmetric hyperbolic system.*

Remark 3.1.4. If Assumption 3.1.4 does not hold then even simple first order deterministic systems with smooth coefficients may be ill-posed. Consider, for example, the system

$$\begin{aligned} du_t(x) =& Dv_t(x)\, dt \\ dv_t(x) =& -Du_t(x)\, dt \end{aligned} \tag{3.1.13}$$

for $(u_t(x), v_t(x))$, $t \in [0, T]$, $x \in \mathbb{R}$, with initial condition $u_0 = \psi$, $v_0 = \phi$, such that $\psi, \phi \in W_2^m \setminus W_2^{m+1}$ for an integer $m \geq 1$. Clearly, this system does not satisfy Assumption 3.1.4, and one can show that it does not have a solution with the initial condition $u_0 = \psi$, $v_0 = \phi$. We note, however, that it is not difficult to show that for any constant $\varepsilon \neq 0$ and Wiener process w the stochastic PDE system

$$\begin{aligned} du_t(x) =& Dv_t(x)\, dt + \varepsilon Dv_t(x)\, dw_t \\ dv_t(x) =& -Du_t(x)\, dt - \varepsilon Du_t(x)\, dw_t \end{aligned} \tag{3.1.14}$$

with initial condition $(u_0, v_0) = (\psi, \phi) \in W_2^m$ (for $m \geq 1$) has a unique solution $(u_t, v_t)_{t \in [0,T]}$, which is a W_2^m-valued continuous process. We leave the proof of this statement and the statement about the nonexistence of a solution to (3.1.13) as exercises for those readers who find that interesting. Clearly, system (3.1.14) does not belong to the class of stochastic systems considered in this paper.

3.2 The main estimate

First let us invoke Itô's formula for the L_p norm. The following in the special case $M = 1$ is Theorem 2.1 from [20]. The proof of this multidimensional variant goes the same way, and therefore will be omitted. Note that for $p \geq 2$ the second derivative, $D_{ij}\langle x\rangle^p$ of the function $(x_1, x_2, \dots, x_M) \to \langle x\rangle^p$ for $p \geq 2$ is

$$p(p-2)\langle x\rangle^{p-4}x_ix_j + p\langle x\rangle^{p-2}\delta_{ij},$$

which makes the last term in (3.2.15) below natural. Here and later on we use the convention $0 \cdot 0^{-1} := 0$ whenever such terms occur.

Lemma 3.2.1. *Let $p \geq 2$ and let $\psi = (\psi^k)_{k=1}^M$ be an $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued $\mathcal{F}_0$-measurable random variable. For $i = 0, 1, 2, ..., d$ and $k = 1, ..., M$ let f^{ki} and $(g^{kr})_{r=1}^\infty$ be predictable functions on $\Omega \times (0, T]$, with values in L_p and in $L_p(l_2)$, respectively, such that*

$$\int_0^T \Big(\sum_{i,k} |f_t^{ki}|_{L_p}^p + \sum_k |g_t^{k\cdot}|_{L_p}^p\Big)\, dt < \infty \quad (a.s.).$$

Suppose that for each $k = 1, ..., M$ we are given a W_p^1-valued predictable function u^k on $\Omega \times (0, T]$ such that

$$\int_0^T |u_t^k|_{W_p^1}^p \, dt < \infty\, (a.s.),$$

and for any $\phi \in C_0^\infty$ with probability 1 for all $t \in [0, T]$ we have

$$(u_t^k, \phi) = (\psi^k, \phi) + \int_0^t (g_s^{kr}, \phi)\, dw_s^r + \int_0^t ((f_s^{k0}, \phi) - (f_s^{ki}, D_i\phi))\, ds.$$

Then there exists a set $\Omega' \subset \Omega$ of full probability such that

$$u = \mathbf{1}_{\Omega'}(u^1, ..., u^k)_{t\in[0,T]}$$

is a continuous $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and for all $t \in [0, T]$

$$\int_{\mathbb{R}^d} \langle u_t\rangle^p \, dx = \int_{\mathbb{R}^d} \langle \psi\rangle^p \, dx + \int_0^t \int_{\mathbb{R}^d} p\langle u_s\rangle^{p-2}\langle u_s, g_s^r\rangle \, dx\, dw_s^r$$

$$+\int_0^t\int_{\mathbb{R}^d}\Big(p\langle u_s\rangle^{p-2}\langle u_s,f^0_s\rangle-p\langle u_s\rangle^{p-2}\langle D_iu_s,f^i_s\rangle$$

$$-(1/2)p(p-2)\langle u_s\rangle^{p-4}\langle u_s,f^i_s\rangle D_i\langle u_s\rangle^2$$

$$+\sum_r\big[(1/2)p(p-2)\langle u_s\rangle^{p-4}\langle u_s,g^r_s\rangle^2+(1/2)p\langle u_s\rangle^{p-2}\langle g^r_s\rangle^2\big]\Big)\,dx\,ds,\qquad(3.2.15)$$

where $f^i:=(f^{ki})_{k=1}^M$ *and* $g^r:=(g^{kr})_{k=1}^M$ *for all* $i=0,1,...,d$ *and* $r=1,2,....$

The following lemma presents the crucial a priori estimate to prove solvability in L_p spaces.

Lemma 3.2.2. *Suppose that Assumptions 3.1.1, 3.1.2, and 3.1.3 hold with* $m\geq 0$. *Assume that* $u=(u_t)_{t\in[0,T]}$ *is a solution of (3.1.1)-(3.1.2) on* $[0,T]$. *Then a.s.* u *is a continuous* $L_p(\mathbb{R}^d,\mathbb{R}^M)$*-valued process, and there is a constant* $N=N(p,K,d,M,K_0)$ *such that*

$$d\int_{\mathbb{R}^d}\langle u_t\rangle^p\,dx+(p/4)\int_{\mathbb{R}^d}\langle u_t\rangle^{p-2}\alpha^{ij}_t\langle D_iu_t,D_ju_t\rangle\,dx\,dt$$

$$\leq p\int_{\mathbb{R}^d}\langle u_t\rangle^{p-2}\langle u_t,\sigma^{ik}D_iu_t+\nu^k_tu_t+g^k_t\rangle\,dx\,dw^k_t$$

$$+N\int_{\mathbb{R}^d}\big[\langle u_t\rangle^p+\langle f_t\rangle^p+\big(\sum_k\langle g^k_t\rangle^2\big)^{p/2}+\big(\sum_k\langle Dg^k_t\rangle^2\big)^{p/2}\big]\,dx\,dt.\qquad(3.2.16)$$

Proof. By Lemma 3.2.1 (a.s.) u is a continuous $L_p(\mathbb{R}^d,\mathbb{R}^M)$-valued process and

$$d\int_{\mathbb{R}^d}\langle u_t\rangle^p\,dx=\int_{\mathbb{R}^d}p\langle u_t\rangle^{p-2}\langle u_t,\sigma^{ik}D_iu_t+\nu^k_tu_t+g^k_t\rangle\,dx\,dw^k_t$$

$$+\int_{\mathbb{R}^d}\Big(p\langle u_t\rangle^{p-2}\langle u_t,b^i_tD_iu_t+c_tu_t+f_t-D_ia^{ij}_tD_ju_t\rangle-p\langle u_t\rangle^{p-2}\langle D_iu_t,a^{ij}_tD_ju_t\rangle$$

$$-(1/2)p(p-2)\langle u_t\rangle^{p-4}D_i\langle u_t\rangle^2\langle u_t,a^{ij}_tD_ju_t\rangle$$

$$+\sum_k\big\{(1/2)p(p-2)\langle u_t\rangle^{p-4}\langle u_t,\sigma^{ik}_tD_iu_t+\nu^k_tu_t+g^k_t\rangle^2$$

$$+(1/2)p\langle u_t\rangle^{p-2}\langle\sigma^{ik}_tD_iu_t+\nu^k_tu_t+g^k_t\rangle^2\big\}\Big)\,dx\,dt.\qquad(3.2.17)$$

Observe that

$$\langle u_t\rangle^{p-2}\langle u_t,f_t\rangle\leq\langle u_t\rangle^p+\langle f_t\rangle^p,\quad\langle u_t\rangle^{p-2}\sum_k\langle g^k_t\rangle^2\leq\langle u_t\rangle^p+\big(\sum_k\langle g^k_t\rangle^2\big)^{p/2},$$

$$\langle u_t\rangle^{p-2}\sum_k\langle\nu^k_tu_t,g^k_t\rangle\leq N\langle u_t\rangle^{p-1}\big(\sum_k\langle g^k_t\rangle^2\big)^{1/2}\leq N\langle u_t\rangle^p+N\big(\sum_k\langle g^k_t\rangle^2\big)^{p/2},$$

$$\langle u_t\rangle^{p-4}\sum_k\langle u_t,g^k_t\rangle^2\le\langle u_t\rangle^{p-2}\sum_k\langle g^k_t\rangle^2\le\langle u_t\rangle^p+\big(\sum_k\langle g^k_t\rangle^2\big)^{p/2},$$

$$\langle u_t\rangle^{p-4}\sum_k\langle u_t,\nu^k_tu_t\rangle\langle u_t,g^k_t\rangle\le N\langle u_t\rangle^{p-1}\big(\sum_k\langle g^k_t\rangle^2\big)^{1/2}\le\langle u_t\rangle^p+\big(\sum_k\langle g^k_t\rangle^2\big)^{p/2},$$

$$\langle u_t\rangle^{p-2}\langle u_t,c_tu_t\rangle\le\langle u_t\rangle^{p-1}\langle c_tu_t\rangle\le|c_t|\langle u_t\rangle^p,$$

where $|c|$ denotes the (Hilbert-Schmidt) norm of c.

This shows how to estimate a few terms on the right in (3.2.17). We write $\xi\sim\eta$ if ξ and η have identical integrals over $\mathbb{R}^d$ and we write $\xi\preceq\eta$ if $\xi\sim\eta+\zeta$ and the integral of ζ over $\mathbb{R}^d$ can be estimated by the coefficient of dt in the right-hand side of (3.2.16). For instance, integrating by parts and using the smoothness of σ^{ik}_t and g^k_t we get

$$p\langle u_t\rangle^{p-2}\langle\sigma^{ik}_tD_iu_t,g^k_t\rangle\preceq-p\sigma^{ik}_t(D_i\langle u_t\rangle^{p-2})\langle u_t,g^k_t\rangle \tag{3.2.18}$$

$$=-p(p-2)\langle u_t\rangle^{p-4}\langle u_t,\sigma^{ik}_tD_iu_t\rangle\langle u_t,g^k_t\rangle,$$

where the first expression comes from the last occurrence of g^k_t in (3.2.17), and the last one with an opposite sign appears in the evaluation of the first term behind the summation over k in (3.2.17). Notice, however, that these calculations are not justified when p is close to 2, since in this case $\langle u_t\rangle^{p-2}$ may not be absolutely continuous with respect to x^i and it is not clear either if $0/0$ should be defined as 0 when it occurs in the second line. For $p=2$ we clearly have $\langle\sigma^{ik}_tD_iu_t,g^k_t\rangle\preceq 0$. For $p>2$ we modify the above calculations by approximating the function $\langle t\rangle^{p-2}$, $t\in\mathbb{R}^M$, by continuously differentiable functions $\phi_n(t)=\varphi_n(\langle t\rangle^2)$ such that

$$\lim_{n\to\infty}\varphi_n(r)=|r|^{(p-2)/2},\quad\lim_{n\to\infty}\varphi_n'(r)=(p-2)\mathrm{sign}(r)|r|^{(p-4)/2}/2$$

for all $r\in\mathbb{R}$, and

$$|\varphi_n(r)|\le N|r|^{(p-2)/2},\quad|\varphi_n'(r)|\le N|r|^{(p-4)/2}$$

for all $r\in\mathbb{R}$ and integers $n\ge1$, where $\varphi_n':=d\varphi_n/dr$ and N is a constant independent of n. Thus instead of (3.2.18) we have

$$p\varphi_n(\langle u_t\rangle^2)\langle\sigma^{ik}_tD_iu_t,g^k_t\rangle\preceq-2p\varphi_n'(\langle u_t\rangle^2)\langle u_t,\sigma^{ik}_tD_iu_t\rangle\langle u_t,g^k_t\rangle, \tag{3.2.19}$$

where

$$|\varphi_n'(\langle u_t\rangle^2)\langle u_t,\sigma^{ik}_tD_iu_t\rangle\langle u_t,g^k_t\rangle|\le N\langle u_t\rangle^{p-2}\langle D_iu_t\rangle\langle g^k_t\rangle \tag{3.2.20}$$

with a constant N independent of n. Letting $n \to \infty$ in (3.2.19) we get

$$p\langle u_t\rangle^{p-2}\langle \sigma_t^{ik} D_i u_t, g_t^k\rangle \preceq -p(p-2)\langle u_t\rangle^{p-4}\langle u_t, \sigma_t^{ik} D_i u_t\rangle\langle u_t, g_t^k\rangle,$$

where, due to (3.2.20), $0/0$ means 0 when it occurs.

These manipulations allow us to take care of the terms containing f and g and show that to prove the lemma we have to prove

$$p(I_0 + I_1 + I_2) + (p/2)I_3 + [p(p-2)/2](I_4 + I_5)$$

$$\preceq -(p/4)\langle u_t\rangle^{p-2}\alpha_t^{ij}\langle D_i u_t, D_j u_t\rangle, \tag{3.2.21}$$

where

$$I_0 = -\langle u_t\rangle^{p-2} D_i a_t^{ij}\langle u_t, D_j u_t\rangle, \quad I_1 = -\langle u_t\rangle^{p-2} a_t^{ij}\langle D_i u_t, D_j u_t\rangle$$

$$I_2 = \langle u_t\rangle^{p-2}\langle u_t, b_t^i D_i u_t\rangle, \quad I_3 = \langle u_t\rangle^{p-2}\sum_k \langle \sigma_t^{ik} D_i u_t + \nu_t^k u_t\rangle^2,$$

$$I_4 = \langle u_t\rangle^{p-4}\sum_k \langle u_t, \sigma_t^{ik} D_i u_t + \nu_t^k u_t\rangle^2, \quad I_5 = -\langle u_t\rangle^{p-4} D_i\langle u_t\rangle^2\langle u_t, a_t^{ij} D_j u_t\rangle.$$

Observe that

$$I_0 = -(1/2)\langle u_t\rangle^{p-2} D_i a_t^{ij} D_j\langle u_t\rangle^2 = -(1/p) D_j\langle u_t\rangle^p D_i a_t^{ij} \preceq 0,$$

by the smoothness of a. Also notice that

$$I_3 \preceq \langle u_t\rangle^{p-2}\sigma_t^{ik}\sigma_t^{jk}\langle D_i u_t, D_j u_t\rangle + I_6,$$

where

$$I_6 = 2\langle u_t\rangle^{p-2}\sigma_t^{ik}\langle D_i u_t, \nu^k u_t\rangle.$$

It follows that

$$pI_1 + (p/2)I_3 \preceq -(p/2)\langle u_t\rangle^{p-2}\alpha_t^{ij}\langle D_i u_t, D_j u_t\rangle + (p/2)I_6.$$

Next,

$$I_4 \preceq \langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk}\langle u_t, D_i u_t\rangle\langle u_t, D_j u_t\rangle + 2\langle u_t\rangle^{p-4}\sigma_t^{ik}\langle u_t, D_i u_t\rangle\langle u_t, \nu_t^k u_t\rangle$$

$$= (1/4)\langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk} D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2 + [2/(p-2)](D_i\langle u_t\rangle^{p-2})\sigma_t^{ik}\langle u_t, \nu_t^k u_t\rangle$$

$$\preceq (1/4)\langle u_t\rangle^{p-4}\sigma_t^{ik}\sigma_t^{jk} D_i\langle u_t\rangle^2 D_j\langle u_t\rangle^2 - [1/(p-2)]I_6 - [2/(p-2)]I_7,$$

where

$$I_7 = \langle u_t \rangle^{p-2} \sigma_t^{ik} \langle u_t, \nu_t^k D_i u_t \rangle.$$

Hence

$$pI_1 + (p/2)I_3 + [p(p-2)/2](I_4 + I_5) \preceq -(p/2)\langle u_t \rangle^{p-2} \alpha_t^{ij} \langle D_i u_t, D_j u_t \rangle$$

$$-[p(p-2)/8]\langle u_t \rangle^{p-4} \alpha_t^{ij} D_i \langle u_t \rangle^2 D_j \langle u_t \rangle^2 - pI_7,$$

and

$$I_2 - I_7 = \langle u_t \rangle^{p-2}(\langle u_t, b_t^i D_i u_t \rangle - \sigma_t^{ik} \langle u_t, \nu_t^k D_i u_t \rangle) = \langle u_t \rangle^{p-2} \langle u_t, \beta_t^i D_i u_t \rangle,$$

with $\beta^i = b^i - \sigma^{ik}\nu^k$. It follows by Remark 3.1.2 that the left-hand side of (3.2.21) is estimated in the order defined by $\preceq$ by

$$-(p/2)\langle u_t \rangle^{p-2} \alpha_t^{ij} \langle D_i u_t, D_j u_t \rangle$$

$$-[p(p-2)/8]\langle u_t \rangle^{p-4} \alpha_t^{ij} D_i \langle u_t \rangle^2 D_j \langle u_t \rangle^2$$

$$+K_0 p \langle u_t \rangle^{p-2} \Big| \sum_{i,j=1}^{d} \alpha_t^{ij} \langle D_i u_t, D_j u_t \rangle \Big|^{1/2} \langle u_t \rangle + h^i D_i \langle u_t \rangle^p$$

$$\preceq -(p/4)\langle u_t \rangle^{p-2} \alpha_t^{ij} \langle D_i u_t, D_j u_t \rangle$$

$$- [p(p-2)/8]\langle u_t \rangle^{p-4} \alpha_t^{ij} D_i \langle u_t \rangle^2 D_j \langle u_t \rangle^2 \rangle, \tag{3.2.22}$$

where the last relation follows from the elementary inequality $ab \le \varepsilon a^2 + \varepsilon^{-1} b^2$. The lemma is proved. $\square$

Remark 3.2.1. In the case that $p = 2$ one can replace condition (3.1.6) with the following: There are constant $K_0, N \ge 0$ such that for all continuously differentiable $\mathbb{R}^M$-valued functions $u = u(x)$ with compact support in $\mathbb{R}^d$ and all values of the arguments we have

$$\int_{\mathbb{R}^d} \langle u, \beta^i D_i u \rangle \, dx \le N \int_{\mathbb{R}^d} \langle u \rangle^2 \, dx$$

$$+ K_0 \int_{\mathbb{R}^d} \big(\Big| \sum_{i,j=1}^{d} \alpha^{ij} \langle D_i u, D_j u \rangle \Big|^{1/2} \langle u \rangle + h^i \langle D_i u, u \rangle \big) \, dx. \tag{3.2.23}$$

This condition is weaker than (3.1.6) as follows from Remark 3.1.2 and still by inspecting the above proof we get that u is a continuous $L_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and there is a constant $N = N(K, d, M, K_0)$ such that (3.2.16) holds with

$p = 2$.

Remark 3.2.2. In the case that $p = 2$ and the magnitudes of the first derivatives of b^i are bounded by K one can further replace condition (3.2.23) with a more tractable one, which is Assumption 3.1.4. Indeed, for $\varepsilon > 0$

$$R := \langle u, (\beta^i - h^i I_M) D_i u \rangle = \tfrac{1}{2}\beta^{ikl} D_i(u^k u^l) + \langle u, (\bar{\beta}^i - h^i I_M) D_i u \rangle$$

$$\leq \tfrac{1}{2}\beta^{ikl} D_i(u^k u^l) + \varepsilon \langle (\bar{\beta}^i - h^i I_M) D_i u \rangle^2/2 + \varepsilon^{-1}\langle u \rangle^2/2.$$

Using Assumption 3.1.4 we get

$$R \leq \tfrac{1}{2}\beta^{ikl} D_i(u^k u^l) + \varepsilon M K_0 \alpha^{ij} \langle D_i u, D_j u \rangle/2 + \varepsilon^{-1}\langle u \rangle^2/2$$

for every $\varepsilon > 0$. Hence by integration by parts we have

$$\int_{\mathbb{R}^d} \langle u, \beta^i D_i u \rangle \, dx \leq N \int_{\mathbb{R}^d} \langle u \rangle^2 \, dx + \int_{\mathbb{R}^d} \langle u, h^i I_M D_i u \rangle \, dx$$

$$+ M K_0 \int_{\mathbb{R}^d} (\varepsilon/2) \alpha^{ij} \langle D_i u_t, D_j u_t \rangle + (\varepsilon^{-1}/2) \langle u \rangle^2 \, dx.$$

Minimising here over $\varepsilon > 0$ we get (3.2.23). In that case again u is a continuous $L_2(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and there is a constant $N = N(K, d, M, K_0)$ such that (3.2.16) holds with $p = 2$.

Remark 3.2.3. If $M = 1$, then condition (3.1.7) is obviously satisfied with $K_0 = 0$ and $h^i = b^i - \sigma^{ik}\nu^k$.

Also note that in the general case, if the coefficients are smoother, then by *formally* differentiating equation (3.1.1) with respect to x^i we obtain a new system of equations for the $M \times d$ matrix-valued function

$$v_t = (v_t^{nm}) = D u_t = (D_m u_t^n).$$

We treat the space of $M \times d$ matrices as a Euclidean Md-dimensional space, the coordinates in which are organized in a special way. The inner product in this space is then just $\langle\langle A, B \rangle\rangle = \mathrm{tr} AB^*$. Naturally, linear operators in this space will be given by matrices like $(T^{(nm)(pj)})$, which transforms an $M \times d$ matrix (A^{pj}) into an $M \times d$ matrix (B^{nm}) by the formula

$$B^{nm} = \sum_{p=1}^{m} \sum_{j=1}^{d} T^{(nm)(pj)} A^{pj}.$$

We claim that the coefficients, the initial value and free terms of the system for v_t satisfy Assumptions 3.1.1, 3.1.2, and 3.1.3 with $m \geq 0$ if Assumptions 3.1.1, 3.1.2, and 3.1.3 are satisfied with $m \geq 1$ for the coefficients, the initial value and free terms of the original system for u_t.

Indeed, as is easy to see, v_t satisfies (3.1.1) with the same σ and a and with $\tilde{b}^i$, $\tilde{c}$, $\tilde{f}$, $\tilde{\nu}^k$, $\tilde{g}^k$ in place of b^i, c, f, ν^k, g^k, respectively, where

$$\tilde{b}^{i(nm)(pj)} = D_m a^{ij}\delta^{pn} + b^{inp}\delta^{jm}, \quad \tilde{c}^{(nm)(pj)} = c^{np}\delta^{mj} + D_m b^{jnp}, \tag{3.2.24}$$

$$\tilde{f}^{nm} = D_m f^n + u^r D_m c^{nr}, \quad \tilde{\nu}^{k(nm)(pj)} = D_m \sigma^{jk}\delta^{np} + \nu^{knp}\delta^{mj},$$
$$\tilde{g}^{knm} = D_m g^{kn} + u^r D_m \nu^{knr}. \tag{3.2.25}$$

Then the left-hand side of the counterpart of (3.1.7) for v is

$$\sum_{m=1}^{d} K_m + \sum_{n=1}^{M} J_n,$$

where (no summation with respect to m)

$$K_m = v^{nm} b^{inr} D_i v^{rm} - \sigma^{ik} v^{nm} \nu^{knr} D_i v^{rm}$$

and (no summation with respect to n)

$$J_n = v^{nm} D_m a^{ij} D_i v^{nj} - \sigma^{ik} v^{nm} D_m \sigma^{jk} D_i v^{nj}.$$

Observe that $D_i v^{nj} = D_{ij} u^n$ implying that

$$\sigma^{ik} D_m \sigma^{jk} D_i v^{nj} = (1/2) D_m(\sigma^{ik}\sigma^{jk}) D_{ij} u^n,$$

$$J_n = (1/2) v^{nm} D_m \alpha^{ij} D_{ij} u^n.$$

By Lemma 1.2.4 for any $\varepsilon > 0$ and n (still no summation with respect to n)

$$J_n \leq N\varepsilon^{-1}\langle\langle v\rangle\rangle^2 + \varepsilon\alpha^{ij} D_{ik} u^n D_{jk} u^n,$$

which along with the fact that $D_{ik} u^n = D_i v^{nk}$ yields

$$\sum_{n=1}^{M} J_n \leq N\varepsilon^{-1}\langle\langle v\rangle\rangle^2 + \varepsilon\alpha^{ij}\langle\langle D_i v, D_j v\rangle\rangle.$$

Upon minimizing with respect to ε we find

$$\sum_{n=1}^{M} J_n \leq N\big(\sum_{i,j=1}^{d} \alpha^{ij}\langle\langle D_i v, D_j v\rangle\rangle\big)^{1/2}\langle\langle v\rangle\rangle.$$

Next, by assumption for any $\varepsilon > 0$ and m (still no summation with respect to m)

$$K_m \leq N\varepsilon^{-1}\langle\langle v\rangle\rangle^2 + \varepsilon\alpha^{ij}D_i v^{rm}D_j v^{rm} + (1/2)h^i D_i \sum_{r=1}^{M}(v^{rm})^2.$$

We conclude as above that

$$\sum_{m=1}^{d} K_m \leq N\big(\sum_{i,j=1}^{d} \alpha^{ij}\langle\langle D_i v, D_j v\rangle\rangle\big)^{1/2}\langle\langle v\rangle\rangle + h^i\langle\langle D_i v, v\rangle\rangle$$

and this proves our claim.

The above calculations show also that the coefficients, the initial value and the free terms of the system for v_t satisfy Assumptions 3.1.1, 3.1.2, and 3.1.4 with $m \geq 0$ if Assumptions 3.1.1, 3.1.2, and 3.1.4 are satisfied with $m \geq 1$ for the coefficients, the initial value and free terms of the original equation for u_t. (Note that due to Assumptions 3.1.1 with $m \geq 1$, $\tilde{b}$, given in (3.2.24), has first order derivatives in x, which in magnitude are bounded by a constant.)

Now higher order derivatives of u are obviously estimated through lower order ones on the basis of this remark without any additional computations. However, we still need to be sure that we can differentiate equation (3.1.1).

Lemma 3.2.3. *Let $m \geq 0$. Suppose that Assumptions 3.1.1, 3.1.2, and 3.1.3 are satisfied and assume that $u = (u_t)_{t\in[0,T]}$ is a solution of* (3.1.1)-(3.1.2) *on* $[0,T]$ *such that (a.s.)*

$$\int_0^T |u_t|^p_{W_p^{m+1}}\,dt < \infty.$$

Then (a.s.) u is a continuous $W_p^m(\mathbb{R}^d,\mathbb{R}^M)$-valued process and for any $q > 0$

$$E\sup_{t\in[0,T]}|u_t|^q_{W_p^m} \leq N(E|\psi|^q_{W_p^m} + E\mathcal{K}^q_{m,p}(T)) \tag{3.2.26}$$

with a constant $N = N(m,p,q,d,M,K,K_0,T)$. If $p = 2$ and instead of Assumption 3.1.3 Assumption 3.1.4 holds and (in case $m = 0$) the magnitudes of the first derivatives of b^i are bounded by K, then u is a continuous $W_2^m(\mathbb{R}^d,\mathbb{R}^M)$-valued process, and for any $q > 0$ estimate (3.2.26) *holds (with $p = 2$).*

Proof. We are going to prove the lemma by induction on m. First let $m = 0$ and denote $y_t := |u_t|^p_{L_p}$. Then by virtue of Remark 3.2.2 and Lemma 3.2.2, the process $y = (y_t)_{t\in[0,T]}$ is an adapted L_p-valued continuous process, and (1.2.2) holds with

$$F_t := \int_{\mathbb{R}^d} \big[\langle f_t\rangle^p + \big(\sum_k \langle g^k_t\rangle^2\big)^{p/2} + \big(\sum_k \langle Dg^k_t\rangle^2\big)^{p/2}\big]\,dx,$$

$$m_t := p\int_0^t \int_{\mathbb{R}^d} \langle u_s\rangle^{p-2}\langle u_s, \sigma^{ik}_s D_i u_s + \nu^k_s u_s + g^k_s\rangle\,dx\,dw^k_s.$$

Notice that

$$d[m_t] = p^2 \sum_{r=1}^{\infty} \Big(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2}\langle u_t, \sigma^{ir}_t D_i u_t + \nu^r_t u_t + g^r_t\rangle\,dx\Big)^2\,dt.$$

$$\le 3p^2(A_t + B_t + C_t)\,dt,$$

with

$$A_t = \sum_{r=1}^{\infty} \Big(p\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2}\sigma^{ir}_t\langle u_t, D_i u_t\rangle\,dx\Big)^2 = \sum_{r=1}^{\infty} \Big(\int_{\mathbb{R}^d} \sigma^{ir}_t D_i\langle u_t\rangle^p\,dx\Big)^2,$$

$$B_t = \sum_{r=1}^{\infty} \Big(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2}\langle u_t, \nu^r_t u_t\rangle\,dx\Big)^2, \quad C_t = \sum_{r=1}^{\infty} \Big(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-2}\langle u_t, g^r_t\rangle\,dx\Big)^2.$$

Integrating by parts and then using Minkowski's inequality, due to Assumption 3.1.1, we get $A_t \le N y^2_t$ with a constant $N = N(K, M, d)$. Using Minkowski's inequality and taking into account that

$$\sum_{r=1}^{\infty} \langle u, \nu^r u\rangle^2 \le \langle u\rangle^4 \sum_{r=1}^{\infty} |\nu^r|^2 \le N\langle u\rangle^4, \quad \sum_{r=1}^{\infty} \langle u, g^r\rangle^2 \le \langle u\rangle^2 |g|,$$

we obtain

$$B_t \le N y^2_t, \quad C_t \le \Big(\int_{\mathbb{R}^d} \langle u_t\rangle^{p-1}|g_t|\,dx\Big)^2 \le |y_t|^{2(p-1)/p}|g_t|^2_{L_p}.$$

Consequently, condition (1.2.3) holds with $G_t = |g_t|^p_{L_p}$, $\rho = 1/p$, and we get (3.2.26) with $m = 0$ by applying Lemma 1.2.5.

Let $m \ge 1$ and assume that the assertions of the lemma are valid for $m-1$, in place of m, for any $M \ge 1$, $p \ge 2$ and $q > 0$, for any u, ψ, f and g satisfying the assumptions with $m-1$ in place of m. Recall the notation $v = (v^{nl}_t) = (D_l u^n_t)$ from Remark 3.2.3, and that v_t satisfies (3.1.1) with the same σ and a and with

$\tilde{b}^i$, $\tilde{c}$, $\tilde{f}$, $\tilde{\nu}^k$, $\tilde{g}^k$ in place of b^i, c, f, ν^k, g^k, respectively. By virtue of Remarks 3.2.3 and 3.2.2 the system for $v = (v_t)_{t\in[0,T]}$ satisfies Assumption 3.1.3, and it is easy to see that it satisfies also Assumptions 3.1.1 and 3.1.2 with $m-1$ in place of m. Hence by the induction hypothesis v is a continuous $W_p^{m-1}(\mathbb{R}^d, \mathbb{R}^M)$-valued adapted process, and we have

$$E \sup_{t\in[0,T]} |v_t|^q_{W_p^{m-1}} \le N(E|\tilde{\psi}|^q_{W_p^{m-1}} + E\tilde{\mathcal{K}}^q_{m-1,p}(T)) \tag{3.2.27}$$

with a constant $N = N(T, K, K_0, M, d, p, q)$, where $\tilde{\psi}^{nl} = D_l\psi^n$,

$$\tilde{\mathcal{K}}^p_{m-1,p}(T) := \int_0^T (|\tilde{f}_t|^p_{W_p^{m-1}} + |\tilde{g}_t|^p_{W_p^m})\, dt.$$

It follows that $(u_t)_{t\in[0,T]}$ is a $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous adapted process, and by using the induction hypothesis it is easy to see that

$$E\tilde{\mathcal{K}}^q_{m-1,p}(T)) \le N(E|\psi|^q_{W_p^m} + E\mathcal{K}^q_{m,p}(T)).$$

Thus (3.2.26) follows.

If $p = 2$ and Assumption 3.1.3 is replaced with Assumptions 3.1.4, then the proof of the conclusion of the lemma goes in the same way with obvious changes. The proof is complete. □

3.3 Proof of the main results

First we prove uniqueness. Let $u^{(1)}$ and $u^{(2)}$ be solutions to (3.1.1)-(3.1.2), and let Assumptions 3.1.1, 3.1.2 and 3.1.3 hold with $m = 0$. Then $u := u^{(1)} - u^{(2)}$ solves (3.1.1) with $u_0 = 0$, $g = 0$ and $f = 0$ and Lemma 3.2.2 and Remark 3.2.2 are applicable to u. Then using Itô's formula for transforming $|u_t|^p_{L_p}\exp(-\lambda t)$ with a sufficiently large constant λ, after simple calculations we get that almost surely

$$0 \le e^{-\lambda t}|u_t|^p_{L_p} \le m_t \quad \text{for all } t \in [0, T],$$

where $m := (m_t)_{t\in[0,T]}$ is a continuous local martingale starting from 0. Hence almost surely $m_t = 0$ for all t, and it follows that almost surely $u^{(1)}_t(x) = u^{(2)}_t(x)$ for all t and almost every $x \in \mathbb{R}^d$. If $p = 2$ and Assumptions 3.1.1, 3.1.2 and 3.1.4 hold and the magnitudes of the first derivatives of b^i are bounded by K and $u^{(1)}$ and $u^{(2)}$ are solutions, then we can repeat the above argument with $p = 2$ to get $u^{(1)} = u^{(2)}$.

To show the existence of solutions we approximate the data of system (3.1.1) with smooth ones, satisfying also the strong stochastic parabolicity.To this end we will use the approximation described in the following lemma.

Lemma 3.3.1. *Let Assumptions 3.1.1 and 3.1.3 (3.1.4, respectively) hold with $m \geq 1$. Then for every $\varepsilon \in (0,1)$ there exist $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable smooth (in x) functions $a^{\varepsilon ij}$, $b^{(\varepsilon)i}$, $c^{(\varepsilon)}$, $\sigma^{(\varepsilon)i}$, $\nu^{(\varepsilon)}, D_k a^{\varepsilon ij}$ and $h^{(\varepsilon)i}$, satisfying the following conditions for every $i,j,k = 1,...,d$.*

(i) There is a constant $N = N(K)$ such that

$$|a^{\varepsilon ij} - a^{ij}| + |b^{(\varepsilon)i} - b^i| + |c^{(\varepsilon)} - c| + |D_k a^{\varepsilon ij} - D_k a^{ij}| \leq N\varepsilon,$$

$$|\sigma^{(\varepsilon)i} - \sigma^i| + |\nu^{(\varepsilon)} - \nu| \leq N\varepsilon$$

for all (ω, t, x) and $i,j,k = 1,...,d$.

(ii) For every integer $n \geq 0$ the partial derivatives in x of $a^{\varepsilon ij}$, $b^{(\varepsilon)i}$, $c^{(\varepsilon)}$, $\sigma^{(\varepsilon)i}$ and $\nu^{(\varepsilon)}$ up to order n are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions, in magnitude bounded by a constant. For $n = m$ this constant is independent of ε, it depends only on m, M, d and K;

(iii) For the matrix $\alpha^{\varepsilon ij} := 2a^{\varepsilon ij} - \sigma^{(\varepsilon)ik}\sigma^{(\varepsilon)jk}$ we have

$$\alpha^{\varepsilon ij}\lambda^i\lambda^j \geq \varepsilon \sum_{i=1}^{d} |\lambda^i|^2 \quad \text{for all } \lambda = (\lambda^1, ..., \lambda^d) \in \mathbb{R}^d;$$

(iv) Assumption 3.1.3 (3.1.4, respectively) holds for the functions $\alpha^{\varepsilon ij}$, $\beta^{\varepsilon i} := b^{(\varepsilon)i} - \sigma^{(\varepsilon)ik}\nu^{(\varepsilon)k}$ and $h^{(\varepsilon)i}$ in place of α^{ij}, β^i and h^i, respectively, with the same constant K_0.

Proof. The proofs of the two statements containing Assumptions 3.1.3 and 3.1.4, respectively, go in essentially the same way, therefore we only detail the former. Let ζ be a nonnegative smooth function on $\mathbb{R}^d$ with unit integral and support in the unit ball, and let $\zeta_\varepsilon(x) = \varepsilon^{-d}\zeta(x/\varepsilon)$. Define

$$b^{(\varepsilon)i} = b^i * \zeta_\varepsilon,\ c^{(\varepsilon)} = c * \zeta_\varepsilon,\ \sigma^{(\varepsilon)i} = \sigma^i * \zeta_\varepsilon,\ \nu^{(\varepsilon)} = \nu * \zeta_\varepsilon,\ h^{(\varepsilon)i} = h^i * \zeta_\varepsilon,$$

and $a^{\varepsilon ij} = a^{ij} * \zeta_\varepsilon + k\varepsilon\delta_{ij}$ with a constant $k > 0$ determined later, where δ_{ij} is the Kronecker symbol and '$*$' means the convolution in the variable $x \in \mathbb{R}^d$. Since we have mollified functions which are bounded and Lipschitz continuous, the mollified functions, together with $a^{\varepsilon ij}$ and $D_k a^{\varepsilon ij}$, satisfy conditions (i) and (ii). Furthermore,

$$|\sigma^{(\varepsilon)ir}\nu^{(\varepsilon)r} - \sigma^{ir}\nu^r| \le |\sigma^{(\varepsilon)i} - \sigma^i||\nu^{(\varepsilon)}| + |\sigma^i||\nu^{(\varepsilon)} - \nu| \le 2K^2\varepsilon,$$

for every $i = 1, ..., d$. Similarly,

$$|\sigma^{(\varepsilon)ir}\sigma^{(\varepsilon)jr} - \sigma^{ir}\sigma^{jr}| \le 2K^2\varepsilon, \quad |b^{(\varepsilon)i} - b^i| \le K\varepsilon, \quad |h^{(\varepsilon)i} - h^i| \le N\varepsilon$$

for all $i, j = 1, 2, ..., d$. Hence setting

$$B^{\varepsilon i} = b^{(\varepsilon)i} - \sigma^{(\varepsilon)ik}\nu^{(\varepsilon)k} - h^{(\varepsilon)i}I_M,$$

and using the notation B^i for the same expression without the superscript 'ε', we have

$$|B^{\varepsilon i} - B^i| \le |b^{(\varepsilon)i} - b^i| + |\sigma^{(\varepsilon)ir}\nu^{(\varepsilon)r} - \sigma^{ir}\nu^r| + \sqrt{M}|h^{(\varepsilon)i} - h^i| \le R\varepsilon,$$

$$|B^{(\varepsilon)i} + B^i| \le R$$

with a constant $R = R(M, K)$. Thus for any z_1,...,z_d vectors from $\mathbb{R}^M$

$$|\langle B^{\varepsilon i}z_i\rangle^2 - \langle B^i z_i\rangle^2| = |\langle (B^{\varepsilon i} - B^i)z_i, (B^{\varepsilon j} + B^j)z_j\rangle|$$

$$\le |B^{\varepsilon i} - B^i||B^{\varepsilon j} + B^j|\langle z_i\rangle\langle z_j\rangle \le dR^2\varepsilon\sum_{i=1}^d \langle z_i\rangle^2.$$

Therefore

$$\langle B^{\varepsilon i}z_i\rangle^2 \le \langle B^i z_i\rangle^2 + C_1\varepsilon\sum_{i=1}^d \langle z_i\rangle^2$$

with a constant $C_1 = C_1(M, K, d)$. Similarly,

$$\sum_{i,j}(2a^{\varepsilon ij} - \sigma^{(\varepsilon)ik}\sigma^{(\varepsilon)jk})\langle z_i, z_j\rangle$$

$$\ge \sum_{i,j}(2a^{ij} - \sigma^{ik}\sigma^{jk})\langle z_i, z_j\rangle + (k - C_2)\varepsilon\sum_i \langle z_i\rangle^2$$

with a constant $C_2 = C_2(K, m, d)$. Consequently,

$$\langle (\beta^{\varepsilon i} - h^{(\varepsilon)i}I_M)z_i\rangle^2 \le \langle B^i z_i\rangle^2 + C_1\varepsilon\sum_{i=1}^d \langle z_i\rangle^2$$

$$\leq K_0 \sum_{i,j=1}^d \alpha^{ij}\langle z_i, z_j\rangle + C_1\varepsilon \sum_{i=1}^d \langle z_i\rangle^2$$

$$\leq K_0 \sum_{i,j=1}^d \alpha^{\varepsilon ij}\langle z_i, z_j\rangle + (K_0(C_2-k)+C_1)\varepsilon \sum_{i=1}^d \langle z_i\rangle^2.$$

Choosing k such that $K_0(C_2-k)+C_1 = -K_0$ we get

$$\langle(\beta^{\varepsilon i} - h^{(\varepsilon)i}I_M)z_i\rangle^2 + K_0\varepsilon \sum_{i=1}^d \langle z_i\rangle^2 \leq K_0 \sum_{i,j=1}^d \alpha^{\varepsilon ij}\langle z_i, z_j\rangle.$$

Hence statements (iii) and (iv) follow immediately. □

Now we start the proof of the existence of solutions which are $W^m_p(\mathbb{R}^d, \mathbb{R}^M)$-valued if the Assumptions 3.1.1, 3.1.2 and 3.1.3 hold with $m \geq 1$. First we make the additional assumptions that ψ, f and g vanish for $|x| \geq R$ for some $R > 0$, and that $q \in [2, \infty)$ and

$$E|\psi|^q_{W^m_p} + E\mathcal{K}^q_{m,q}(T) < \infty. \tag{3.3.28}$$

For each $\varepsilon > 0$ we consider the system

$$du^\varepsilon_t = [\sigma^{(\varepsilon)ir}_t D_i u^\varepsilon_t + \nu^{(\varepsilon)r}_t u^\varepsilon_t + g^{(\varepsilon)r}_t]\, dw^r_t$$

$$+ \left[a^{\varepsilon ij}_t D_{ij} u^\varepsilon_t + b^{(\varepsilon)i}_t D_i u^\varepsilon_t + f^{(\varepsilon)}_t\right] dt \tag{3.3.29}$$

with initial condition

$$u^{(\varepsilon)}_0 = \psi^{(\varepsilon)}, \tag{3.3.30}$$

where the coefficients are taken from Lemma 3.3.1, and $\psi^{(\epsilon)}$, $f^{(\epsilon)}$ and $g^{(\epsilon)}$ are defined as the convolution of ψ, f and g, respectively, with $\zeta_\varepsilon(\cdot) = \varepsilon^{-d}\zeta(\cdot/\varepsilon)$ for $\zeta \in C^\infty_0(\mathbb{R}^d)$ taken from the proof of Lemma 3.3.1. By Theorem 1.0.1 the above equation has a unique solution u^ε, which is a $W^n_2(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous process for all n. Hence, by Sobolev embeddings, u^ε is a $W^{m+1}_p(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous process, and therefore we can use Lemma 3.2.3 to get

$$E \sup_{t\in[0,T]} |u^\varepsilon_t|^q_{W^n_{p'}} \leq N(E|\psi^{(\varepsilon)}|^q_{W^n_{p'}} + E(\mathcal{K}^\varepsilon_{n,p'})^q(T)) \tag{3.3.31}$$

for $p' \in \{p, 2\}$ and $n = 0, 1, 2...m$, where $K^\varepsilon_{n,p'}$ is defined by (3.1.4) with $f^{(\varepsilon)}$ and $g^{(\varepsilon)}$ in place of f and g, respectively. Keeping in mind that $T^{1/r} \leq \max\{1, T\}$, and using basic properties of convolution, we can conclude that

$$E\left(\int_0^T |u_t^\varepsilon|_{W_{p'}^n}^r\,dt\right)^{q/r} \le N(E|\psi|_{W_{p'}^n}^q + E\mathcal{K}_{n,p'}^q(T)) \tag{3.3.32}$$

for any $r > 1$ and with $N = N(m,p,q,d,M,K,T)$ not depending on r.

For integers $n \ge 0$, and any $r, q \in (1,\infty)$ let $\mathbb{H}_{p,r,q}^n$ be the space of $\mathbb{R}^M$-valued functions $v = v_t(x) = (v_t^i(x))_{i=1}^M$ on $\Omega \times [0,T] \times \mathbb{R}^d$ such that $v = (v_t(\cdot))_{t\in[0,T]}$ are $W_p^n(\mathbb{R}^d,\mathbb{R}^M)$-valued predictable processes and

$$|v|_{\mathbb{H}_{p,r,q}^n}^q = E\left(\int_0^T |v_t|_{W_p^n}^r\,dt\right)^{q/r} < \infty.$$

Then $\mathbb{H}_{p,r,q}^n$ with the norm defined above is a reflexive Banach space for each $n \ge 0$ and $p, r, q \in (1,\infty)$. We use the notation $\mathbb{H}_{p,q}^n$ for $\mathbb{H}_{p,q,q}^n$.

By Assumption 3.1.2 the right-hand side of (3.3.32) is finite for $p' = p$ and also for $p = 2$ since ψ, f and g vanish for $|x| \ge R$. Thus there exists a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ such that $\epsilon_k \to 0$ and for $p' = p, 2$ and integers $r > 1$ and $n \in [0,m]$ the sequence $v^k := u^{\varepsilon_k}$ converges weakly in $\mathbb{H}_{p'r,q}^n$ to some $v \in H_{p',r,q}^m$, which therefore also satisfies

$$E\left(\int_0^T |v_t|_{W_{p'}^n}^r\,dt\right)^{q/r} \le N(E|\psi|_{W_{p'}^n}^q + E\mathcal{K}_{n,q}^q(T))$$

for $p' = p, 2$ and integers $r > 1$. Using this with $p' = p$ and letting $r \to \infty$ by Fatou's lemma we obtain

$$E\text{esssup}_{t\in[0,T]}|v_t|_{W_p^n}^q \le N(E|\psi|_{W_p^n}^q + E\mathcal{K}_{n,p}^q(T)) \quad \text{for } n = 0,1,...,m. \tag{3.3.33}$$

Now we are going to show that a suitable stochastic modification of v is a solution of (3.1.1)-(3.1.2). To this end we fix an $\mathbb{R}^M$-valued function φ in $C_0^\infty(\mathbb{R}^d)$ and a predictable real-valued process $(\eta_t)_{t\in[0,T]}$, which is bounded by some constant C, and define the functionals Φ, Φ_k, Ψ and Ψ_k over $\mathbb{H}_{p,q}^1$ by

$$\Phi_k(u) = E\int_0^T \eta_t \int_0^t \{-(a_s^{\varepsilon_k ij} D_i u_s, D_j\varphi) + (\bar{b}_s^{\varepsilon_k i} D_i u_s + c_s^{(\varepsilon_k)} u_s, \varphi)\}\,ds\,dt,$$

$$\Phi(u) = E\int_0^T \eta_t \int_0^t \{-(a_s^{ij} D_i u_s, D_j\varphi) + (\bar{b}_s^i D_i u_s + c_s u_s, \varphi)\}\,ds\,dt,$$

$$\Psi(u) = E\int_0^T \eta_t \int_0^t (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)\,dw_t^r\,dt$$

$$\Psi_k(u) = E\int_0^T \eta_t \int_0^t (\sigma_t^{(\varepsilon_k)ir} D_i u_t + \nu_t^{(\varepsilon_k)r} u_t, \varphi)\, dw_t^r\, dt$$

for $u \in \mathbb{H}^1_{p,q}$ for each $k \geq 1$, where $\bar{b}^{\varepsilon i} = b^{(\varepsilon)i} - D_j a^{\varepsilon ij} I_M$. By the Bunyakovsky-Cauchy-Schwarz and the Burkholder-Davis-Gundy inequalities for all $u \in \mathbb{H}^1_{p,q}$ we have

$$\Phi(u) \leq CNT^{2-1/q} |u|_{\mathbb{H}^1_{p,q}} |\varphi|_{W^1_{\bar{p}}},$$

$$\begin{aligned}
\Psi(u) \leq & CTE \sup_{t\leq T} \left| \int_0^t (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)\, dw_t^r \right| \\
\leq & 3CTE \left\{ \int_0^T \sum_{r=1}^{\infty} (\sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi)^2\, dt \right\}^{1/2} \\
\leq & 3CTE \left\{ \int_0^T \left(\int_{\mathbb{R}^d} |\langle \sigma_t^{ir} D_i u_t + \nu_t^r u_t, \varphi \rangle|_{l_2}\, dx \right)^2 dt \right\}^{1/2} \\
\leq & CTNE \left\{ \int_0^T |u_t|^2_{W^1_p} |\varphi|^2_{W^1_{\bar{p}}}\, dt \right\}^{1/2} \leq CNT^{q/2} |u|_{\mathbb{H}^1_{p,q}} |\varphi|_{W^1_{\bar{p}}}
\end{aligned}$$

with a constant $N = N(K, d, M)$, where $\bar{p} = p/(p-1)$. (In the last inequality we make use of the assumption $q \geq 2$.) Consequently, Φ and Ψ are continuous linear functionals over $\mathbb{H}^1_{p,q}$, and therefore

$$\lim_{k\to\infty} \Phi(v^k) = \Phi(v), \quad \lim_{k\to\infty} \Psi(v^k) = \Psi(v). \tag{3.3.34}$$

Using statement (i) of Lemma 3.3.1, we get

$$|\Phi_k(u) - \Phi(u)| + |\Psi_k(u) - \Psi(u)| \leq N\varepsilon_k |u|_{\mathbb{H}^1_{p,q}} |\varphi|_{W^1_{\bar{p}}} \tag{3.3.35}$$

for all $u \in \mathbb{H}^1_{p,q}$ with a constant $N = N(k, d, M)$. Since u^ε is the solution of (3.3.29)-(3.3.30), we have

$$\begin{aligned}
E\int_0^T \eta_t (v_t^k, \varphi)\, dt = & E\int_0^T \eta_t (\psi^k, \varphi)\, dt + \Phi(v^k) + \Psi(v^k) \\
& + F(f^{(\varepsilon_k)}) + G(g^{(\varepsilon_k)})
\end{aligned} \tag{3.3.36}$$

for each k, where

$$F(f^{(\varepsilon_k)}) = E\int_0^T \eta_t \int_0^t (f_s^{(\varepsilon_k)}, \varphi)\, ds\, dt,$$

$$G(g^{(\varepsilon_k)}) = E\int_0^T \eta_t \int_0^t (g_s^{(\varepsilon_k)r}, \varphi)\, dw_s^r\, dt.$$

Taking into account that $|v^k|_{\mathbb{H}^1_{p,q}}$ is a bounded sequence, from (3.3.34) and (3.3.35) we obtain

$$\lim_{k\to\infty} \Phi_n(v^k) = \Phi(v), \quad \lim_{k\to\infty} \Psi_k(v^k) = \Psi(v). \tag{3.3.37}$$

One can see similarly (in fact easier), that

$$\lim_{k\to\infty} E\int_0^T \eta_t(v_t^k, \varphi)\, dt = E\int_0^T \eta_t(v_t, \varphi)\, dt, \tag{3.3.38}$$

$$\lim_{k\to\infty} E\int_0^T \eta_t(\psi_t^{(\varepsilon_k)}, \varphi)\, dt = E\int_0^T \eta_t(\psi, \varphi)\, dt, \tag{3.3.39}$$

$$\lim_{k\to\infty} F(f^{(\varepsilon_k)}) = F(f), \quad \lim_{k\to\infty} G(g^{(\varepsilon_k)}) = G(g). \tag{3.3.40}$$

Letting $k \to \infty$ in (3.3.36), and using (3.3.37) through (3.3.40) we obtain

$$E\int_0^T \eta_t(v_t, \varphi)\, dt$$

$$= E\int_0^T \eta_t \Big\{ (\psi, \varphi) + \int_0^t \big[-(a_s^{ij} D_i u_s, D_j \varphi) + (\bar{b}_s^i D_i u_s + c_s u_s + f_s, \varphi) \big]\, ds$$
$$+ \int_0^t (\sigma^{ir} D_i v_s + \nu^r v_s, \varphi)\, dw_s^r \Big\}\, dt$$

for every bounded predictable process $(\eta_t)_{t\in[0,T]}$ and φ from C_0^∞. Hence for each $\varphi \in C_0^\infty$

$$(v_t, \varphi) = (\psi, \varphi) + \int_0^t \big[-(a_s^{ij} D_i v_s, D_j \varphi) + (\bar{b}_s^i D_i v_s + c_s v_s + f_s, \varphi) \big]\, ds$$

$$+ \int_0^t (\sigma^{ir} D_i v_s + \nu^r v_s + g_s^r, \varphi)\, dw_s^r$$

holds for $P \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$. Substituting here $(-1)^{|\alpha|} D^\alpha \varphi$ in place of φ for a multi-index $\alpha = (\alpha_1, ..., \alpha_d)$ of length $|\alpha| \le m - 1$ and integrating by parts, we see that

$$(D^\alpha v_t, \varphi) = (D^\alpha \psi, \varphi) + \int_0^t \big[-(F_s^j, D_j \varphi) + (F_s^0, \varphi) \big]\, ds + \int_0^t (G_s^r, \varphi)\, dw_s^r \tag{3.3.41}$$

for $P \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$, where, owing to the fact that (3.3.33) also holds with 2 in place of p, F^i and $(G^r)_{r=1}^\infty$ are predictable processes with values in L_2-spaces for $i = 0, 1, ..., d$, such that

$$\int_0^T \big(\sum_{i=0}^d |F_s^i|_{L_2}^2 + |G_s|_{L_2}^2 \big) \, ds < \infty \quad \text{(a.s.)}.$$

Hence the theorem on Itô's formula from [25] implies that in the equivalence class of v in $\mathbb{H}_{2,q}^m$ there is a $W_2^{m-1}(\mathbb{R}^d, \mathbb{R}^M)$-valued continuous process, $u = (u_t)_{t\in[0,T]}$, and (3.3.41) with u in place of v holds for any $\varphi \in C_0^\infty(\mathbb{R}^d)$ almost surely for all $t \in [0, T]$. After that an application of Lemma 3.2.1 to $D^\alpha u$ for $|\alpha| \le m - 1$ yields that $D^\alpha u$ is an $L_p(\mathbb{R}^d, \mathbb{R}^M)$-valued, strongly continuous process for every $|\alpha| \le m - 1$, i.e., u is a $W_p^{m-1}(\mathbb{R}^d, \mathbb{R}^M)$-valued strongly continuous process. This, (3.3.33), and the denseness of C_0^∞ in $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$ implies that (a.s.) u is a $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$-valued weakly continuous process and (3.1.11) holds.

To prove the theorem without the assumption that ψ, f and g have compact support, we take a $\zeta \in C_0^\infty(\mathbb{R}^d)$ such that $\zeta(x) = 1$ for $|x| \le 1$ and $\zeta(x) = 0$ for $|x| \ge 2$, and define $\zeta_n(\cdot) = \zeta(\cdot/n)$ for $n > 0$. Let $u(n) = (u_t(n))_{t\in[0,T]}$ denote the solution of (3.1.1)-(3.1.2) with $\zeta_n\psi$, $\zeta_n f$ and $\zeta_n g$ in place of ψ, f and g, respectively. By virtue of what we have proved above, $u(n)$ is a weakly continuous $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$-valued process, and

$$E \sup_{t\in[0,T]} |u_t(n) - u_t(l)|_{W_p^m}^q \le NE|(\zeta_n - \zeta_l)\psi|_{W_p^m}^q$$

$$+NE\big(\int_0^T \{ |(\zeta_n - \zeta_l) f_s|_{W_p^m}^p + |(\zeta_n - \zeta_l) g_s|_{W_p^{m+1}}^p \} \, ds \big)^{q/p}.$$

Letting here $n, l \to \infty$ and applying Lebesgue's theorem on dominated convergence in the left-hand side, we see that the right-hand side of the inequality tends to zero. Thus for a subsequence $n_k \to \infty$ we have that $u_t(n_k)$ converges strongly in $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$, uniformly in $t \in [0, T]$, to a process u. Hence u is a weakly continuous $W_p^m(\mathbb{R}^d, \mathbb{R}^M)$-valued process. It is easy to show that it solves (3.1.1)-(3.1.2) and satisfies (3.1.11).

By using a standard stopping time argument we can dispense with condition (3.3.28). Finally we can prove estimate (3.1.11) for $q \in (0, 2)$ by applying Lemma 1.2.3 in the usual way. The proof of the Theorem 3.1.1 is complete. We have already showed the uniqueness statement of Theorem 3.1.2, the proof of the other assertions goes in the above way with obvious changes.

Chapter 4

Degenerate equations - numerics

As mentioned in the previous chapter, the study of degenerate equations is motivated by practical applications such as the nonlinear filtering problem and therefore numerical methods to approximate the solution are of interest. However, many approximation results for SPDEs rely strongly on the strong parabolicity. Here we discuss finite difference approximations, motivated by [13], due to four important favourable properties: 1) Easy implementation 2) Availability of pointwise convergence 3) Enough flexibility to cover degenerate equations 4) Expansion of the error to a power series. The latter one is particularly useful when combined with the classical idea of Richardson's extrapolation from [36], to obtain higher order schemes. Such an acceleration of the convergence of the spatial discretization is established in [13]. Below we attempt, with partial success, to relax the smoothness conditions on the coefficients. Also, we discuss the error one makes when they solve a truncated version of the equation, which is a necessary but rarely discussed step to make the implementation of the scheme feasible. We apply this error estimate, along with the results of [13] and the analysis of the implicit Euler method for degenerate equations, to obtain a fully discrete, implementable scheme. The content of this chapter is based on the papers [9], [10], joint works with István Gyöngy.

4.1 L_p estimates and acceleration - Formulation

We consider the SPDE

$$du_t(x) = [D_i(a_t^{ij}(x)D_ju_t(x)) + b_t^i(x)D_iu_t(x) + c_t(x)u_t(x) + f_t(x)]\,dt$$

$$+ [\sigma_t^{ir}D_iu_t(x) + \nu_t^r(x)u_t(x) + g^r(x)]\,dw_t^r \tag{4.1.1}$$

for $(t, x) \in [0, T] \times \mathbb{R}^d$, with the initial condition

$$u_0(x) = \psi(x) \quad x \in \mathbb{R}^d. \tag{4.1.2}$$

Because of the different form of the equation and because some different notations will be more convenient, we formulate the similar assumptions to Chapter 3 again, for integers $m \geq 1$.

To introduce the finite difference schemes approximating (4.1.1) first let Λ_0, $\Lambda_1 \subset \mathbb{R}^d$ be two finite sets, the latter being symmetric to the origin, and $0 \in \Lambda_1 \setminus \Lambda_0$. Denote

$$\Lambda = \Lambda_0 \cup -\Lambda_0 \cup \Lambda_1$$

and $|\Lambda| = \sum_{\lambda \in \Lambda} |\lambda|$. On Λ we make the following natural assumption: If any subset $\Lambda' \subset \Lambda$ is linearly dependent, then Λ' is linearly dependent over the rationals. This ensures that the following grid is locally finite. Let $\mathbb{G}_h$ denote the grid

$$\mathbb{G}_h = \{h(\lambda_1 + \ldots + \lambda_n) : \lambda_i \in \Lambda, n = 1, 2, ...\},$$

for $h > 0$, and define the finite difference operators

$$\delta_{h,\lambda}\varphi(x) = (1/h)(\varphi(x + h\lambda) - \varphi(x))$$

and the shift operators

$$T_{h,\lambda}\varphi(x) = \varphi(x + h\lambda)$$

for $\lambda \in \Lambda$ and $h \neq 0$. Notice that $\delta_{h,0}\varphi = 0$ and $T_{h,0}\varphi = \varphi$. For a fixed $h > 0$ consider the finite difference equation

$$du_t^h(x) = (L_t^h(x)u_t^h(x) + f_t(x))\, dt + (M_t^{hr}(x)u_t^h(x) + g_t^r(x))\, dw_t^r, \tag{4.1.3}$$

for $(t, x) \in [0, T] \times \mathbb{G}_h$, with the initial condition

$$u_0^h(x) = \psi(x) \tag{4.1.4}$$

for $x \in \mathbb{G}_h$, where

$$L_t^h\varphi = \sum_{\lambda \in \Lambda_0} \delta_{-h,\lambda}(\mathfrak{a}_h^\lambda \delta_{h,\lambda}\varphi) + \sum_{\gamma \in \Lambda_1} \mathfrak{p}_h^\gamma \delta_{h,\gamma}\varphi + \sum_{\gamma \in \Lambda_1} \mathfrak{c}_h^\gamma T_{h,\gamma}\varphi$$

and

$$M_t^{hr}\varphi = \sum_{\lambda \in \Lambda_0} \mathfrak{s}_h^{\lambda r} \delta_{h,\lambda}\varphi + \sum_{\gamma \in \Lambda_1} \mathfrak{n}_h^{\gamma r} T_{h,\gamma}\varphi$$

for functions φ on $\mathbb{G}_h$. The coefficients $\mathfrak{a}_h^\lambda$, $\mathfrak{p}_h^\gamma$, and $\mathfrak{c}_h^\gamma$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable bounded functions on $\Omega \times [0,T] \times \mathbb{R}^d$, with values in $\mathbb{R}$, and $\mathfrak{p}_h^0 = 0$ is assumed. The coefficients $\mathfrak{s}_h^\lambda$ and $\mathfrak{n}_h^\gamma$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable bounded functions on $\Omega \times [0,T] \times \mathbb{R}^d$, with values in l_2. All of them are supposed to be defined for $h = 0$ as well, and to depend continuously on h.

One can look for solutions of the above scheme in the space of adapted stochastic processes with values in $l_{p,h}$, the space of real functions ϕ on $\mathbb{G}_h$ such that

$$|\phi|_{l_{p,h}}^p = \sum_{x \in \mathbb{G}_h} |\phi(x)|^p h^d < \infty.$$

The similar space is defined for l_2-valued functions and will be denoted by $l_{p,h}(l_2)$. For a fixed h equation (4.1.3) is an SDE in $l_{p,h}$, with Lipschitz coefficients. Hence if almost surely

$$|\psi|_{l_{p,h}}^p + \int_0^T |f_t|_{l_{p,h}}^p + |g_t|_{l_{p,h}(l_2)}^p \, dt < \infty,$$

then (4.1.3)-(4.1.4) admits a unique $l_{p,h}$-valued solution $(u_t^h)_{t \in [0,T]}$.

Remark 4.1.1. By well-known results on Sobolev embeddings, if $m > k + d/p$, there exists a bounded operator J from W_p^m to the space of functions with bounded and continuous derivatives up to order k such that $Jv = v$ almost everywhere. We will always identify functions with their continuous modifications if they have one, without introducing new notation for them. It is also known, and can be easily seen, that if $m > d/p$, then the for $v \in W_p^m$ the restriction of Jv onto the grid $\mathbb{G}_h$ is in $l_{p,h}$, moreover,

$$|Jv|_{l_{p,h}} \leq C|v|_{W_p^m}, \tag{4.1.5}$$

where C is independent of v and h.

Remark 4.1.2. The h-dependency of the coefficients may seem artificial and in fact does not mean any additional difficulty in the proof of Theorems 4.1.1-4.1.3 below. However, we will make use of this generality to extend our results to the case when the data in the problem (4.1.1)-(4.1.2) are in some weighted Sobolev spaces.

Clearly

$$\delta_{h,\lambda}\varphi(x) \to \partial_\lambda \varphi(x)$$

as $h \to 0$ for smooth functions φ, so in order to get that our finite difference operators approximate the corresponding differential operators, we make the following assumption.

Assumption 4.1.1. We have, for every $i, j = 1, \dots, d$, $r = 1, \dots$

$$a^{ij} = \sum_{\lambda \in \Lambda_0} \mathfrak{a}_0^\lambda \lambda^i \lambda^j, \tag{4.1.6}$$

$$b^i = \sum_{\gamma \in \Lambda_1} \mathfrak{p}_0^\gamma \gamma^i, \ c = \sum_{\gamma \in \Lambda_1} \mathfrak{c}_0^\gamma, \tag{4.1.7}$$

$$\sigma^{ir} = \sum_{\lambda \in \Lambda_0} \mathfrak{s}_0^{\lambda r} \lambda^i, \ \nu^r = \sum_{\gamma \in \Lambda_1} \mathfrak{n}_0^{\gamma r} \tag{4.1.8}$$

and for $P \times dt \times dx$-almost all (ω, t, x) we have for all $(z_\lambda)_{\lambda \in \Lambda_0}$

$$\mathfrak{a}_h^\lambda (z_\lambda)^2 - 2p \mathfrak{s}_h^{\lambda r} \mathfrak{s}_h^{\mu r} z_\lambda z_\mu \geq 0, \ \mathfrak{p}_h^\gamma \geq 0 \quad \text{for every } \gamma \in \Lambda_1, \, h \geq 0. \tag{4.1.9}$$

Remark 4.1.3. The restriction (4.1.6) together with $\mathfrak{a}_0^\lambda \geq 0$ is not too severe, we refer the reader to [24] for a detailed discussion about matrix-valued functions which possess this property.

Remark 4.1.4. The parabolicity condition in (4.1.9) depends on p. This is an essential restriction, but for example, additive and multiplicative noises are still covered. It is worth mentioning that while unusual, there exist problems where the stochastic parabolicity condition has to depend on p, see e.g. [1]. It is unclear whether this is one of them or our condition can be significantly improved.

Example 4.1.1. Suppose that the matrix (a^{ij}) is diagonal. Then taking $\Lambda_0 = \{e_i : i = 1 \dots d\}$ and $\Lambda_1 = \{0\} \cup \{\pm e_i : i = 1 \dots d\}$, where (e_i) is the standard basis in $\mathbb{R}^d$, one can set

$$\mathfrak{a}_h^{e_i} = a^{ii}, \, \mathfrak{p}_h^{e_i} = b^i + \theta^i, \, \mathfrak{p}_h^{-e_i} = \theta_i, \, \mathfrak{c}_h^0 = c, \, \mathfrak{p}_h^0 = \mathfrak{c}_h^{\pm e_i} = 0,$$

$$\mathfrak{s}_h^{e_i r} = \sigma^{ir}, \, \mathfrak{n}_h^{0r} = \nu^r, \, \mathfrak{n}_h^{\pm e_i r} = 0,$$

with any $\theta^i \geq \max(0, -b^i)$, $i = 1 \dots d$.

Example 4.1.2. Suppose that (a^{ij}) is a $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable function of (ω, t, x) with values in a closed bounded polyhedron in the set of symmetric non-negative $d \times d$ matrices, such that its first and second order derivatives in $x \in \mathbb{R}^d$ are continuous in x and are bounded by a constant K. Then it is shown in [24] that one can obtain a finite set $\Lambda_0 \subset \mathbb{R}^d$ and $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable, bounded, nonnegative functions $\mathfrak{a}_0^\lambda$, $\lambda \in \Lambda_0$ such that (4.1.6) holds, and the first order derivatives of $(\mathfrak{a}_0^\lambda)^{1/2}$ in x are bounded by a constant N depending only on K, d and the polyhedron. Such situation arises in applications when, for example,

$(a^{ij}_t(x))$ is a diagonally dominant symmetric non-negative definite matrix for each (ω, t, x), which by definition means that

$$2a^{ii}_t(x) \geq \sum_{j=1}^{d} |a^{ij}_t(x)|, \quad \text{for all } i = 1, 2, .., d, \text{ and } (\omega, t, x),$$

and hence it clearly follows that (a^{ij}) takes values in a closed polyhedron in the set of symmetric non-negative $d \times d$ matrices. Clearly, this polyhedron can be chosen to be bounded if (a^{ij}) is a bounded function.

Since the compatibility condition (4.1.6)-(4.1.7) will always be assumed, any subsequent conditions will be formulated for the coefficients in (4.1.3), which then automatically imply the corresponding properties for the coefficients in (4.1.1).

Assumption 4.1.2. The derivatives in (h, x) of $\mathfrak{a}^{\lambda}_h$, $\mathfrak{s}^{\lambda}_h$, $\mathfrak{n}^{\gamma}_h$ (resp., $\mathfrak{p}^{\gamma}_h$, $\mathfrak{c}^{\gamma}_h$), up to order $m+1$ (resp., m) are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions bounded by K.

Assumption 4.1.3. The free data, $(f_t)_{t\in[0,T]}$ and $(g_t)_{t\in[0,T]}$ are predictable processes with values in W^m_p and $W^{m+1}_p(\mathbb{R}^d, l_2)$, respectively, such that almost surely

$$F_{m,p}(T) + G_{m,p}(T) := \left(\int_0^T |f_t|^p_{W^m_p} dt\right)^{1/p} + \left(\int_0^T |g_t|^p_{W^m_p} dt\right)^{1/p} < \infty.$$

The initial value, ψ is an $\mathcal{F}_0$-measurable W^m_p-valued random variable.

We are now about to present the main results. The first three theorems correspond to similar results in the L_2 setting from [13]. The key role in their proof is played by Theorem 4.2.3 below, which presents an upper bound for the W^m_p norms of the solutions to (4.1.3)-(4.1.4). After obtaining this estimate, Theorems 4.1.1 through 4.1.3 can be proved in the same fashion as their counterparts in the L_2 setting, therefore, only a sketch of the proof will be provided in which we highlight the main differences; for the complete argument we refer to [13].

Theorem 4.1.1. *Let $k \geq 0$ be an integer and let Assumptions 4.1.1 through 4.1.3 hold with $m > 2k+3+d/p$. Then there are continuous random fields $u^{(1)}, \ldots u^{(k)}$ on $[0,T] \times \mathbb{R}^d$, independent of h, such that almost surely*

$$u^h_t(x) = \sum_{j=0}^{k} \frac{h^j}{j!} u^{(j)}_t(x) + h^{k+1} r^h_t(x) \tag{4.1.10}$$

for $t \in [0,T]$ and $x \in \mathbb{G}_h$, where $u^{(0)} = u$, r^h is a continuous random field on

$[0,T]\times\mathbb{R}^d$, *and for any* $q>0$

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|r^h_t(x)|^q+E\sup_{t\in[0,T]}|r^h_t|^q_{l_{p,h}}\le N(E|\psi|^q_{W^m_p}+EF^q_{m,p}(T)+EG^q_{m,p}(T))$$

with $N=N(K,T,m,p,q,d,|\Lambda|)$.

Once we have the expansion above, we can use Richardson extrapolation to improve the rate of convergence. For a given k set

$$(c_0,c_1,\dots,c_k)=(1,0,0,\dots,0)V^{-1}, \tag{4.1.11}$$

where V denotes the $(k+1)\times(k+1)$ Vandermonde matrix $V=(V^{ij})=(2^{-(i-1)(j-1)})$, and define

$$v^h=\sum_{i=0}^{k}c_iu^{h_i},$$

where $h_i=h/2^i$.

Theorem 4.1.2. *Let* $k\ge 0$ *be an integer and let Assumptions 4.1.1 through 4.1.3 hold with* $m>2k+3+d/p$. *Then for every* $q>0$ *we have*

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|u_t(x)-v^h_t(x)|^q+E\sup_{t\in[0,T]}|u_t-v^h_t|^q_{l_{p,h}}$$

$$\le h^{q(k+1)}N(E|\psi|^q_{W^m_p}+EF^q_{m,p}(T)+EG^q_{m,p}(T))$$

with $N=N(K,T,m,k,p,q,d,|\Lambda|)$.

Theorem 4.1.3. *Let* $(h_n)_{n=1}^{\infty}\in l_q$ *be a nonnegative sequence for some* $q\ge 1$. *Let* $k\ge 0$ *be an integer and let Assumptions 4.1.1 through 4.1.3 hold with* $m>2k+3+d/p$. *Then for every* $\varepsilon>0$ *there exists a random variable* ξ_ε *such that almost surely*

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|u_t(x)-v^h_t(x)|\le\xi_\varepsilon h^{k+1-\varepsilon}$$

for $h=h_n$.

Remark 4.1.5. We can use $h_i=h/n_i$, $i=1\dots k$, with any set of different integers n_i, with $n_1=1$. Then changing the matrix V to $\tilde V=(\tilde V^{ij})=(n_i^{-j+1})$ in (4.1.11), Theorems 4.1.2-4.1.3 remain valid. The choice $n_i=i$, for example, yields a more coarse grid, and can reduce computation time.

Choosing p large enough, in some cases one can get rid of the term d/p in the conditions of the theorems above, thus obtaining dimension-invariant conditions.

To this end, first denote the function $\rho_s(x) = 1/(1+|x|^2)^{s/2}$ defined on $\mathbb{R}^d$ for all $s \geq 0$. We say that a function F on $\mathbb{R}^d$ has polynomial growth of order s if the L_∞ norm of $F\rho_s$ is finite. For any integer $m \geq 0$, the set of functions on $\mathbb{R}^d$ which have polynomial growth of order s and whose derivatives up to order m are functions and have polynomial growth of order s is denoted by P_s^m, and its equipped with the norm

$$\|F\|_{P_s^m} = |F\rho_s|_{W_\infty^m} < \infty.$$

The similar space is defined for l_2-valued functions and is denoted by $P_s^m(l_2)$. Note that for any integers $m > k \geq 0$, if $F \in P_s^m$, then its partial derivatives up to order k exist in the classical sense and along with F are continuous functions. The polynomial growth property of order s for functions on $\mathbb{G}_h$ can also be defined analogously, the set of such functions is denoted by $P_{h,s}$.

Let $s \geq 0$ and m be a nonnegative integer. Consider again the equation

$$\begin{aligned} du_t(x) = (D_i(a_t^{ij}(x)D_ju_t(x)) + b_t^i(x)D_ju_t(x) + c_t(x)u_t(x) + f_t(x))\,dt \\ + (\sigma_t^{ir}(x)D_iu_t(x) + \nu_t^r(x)u_t(x) + g^r(x))\,dw_t^r \end{aligned} \tag{4.1.12}$$

for $(t,x) \in [0,T] \times \mathbb{R}^d$, with the initial condition

$$u_0(x) = \psi(x) \quad x \in \mathbb{R}^d, \tag{4.1.13}$$

where we keep all our measurability conditions from (4.1.1)-(4.1.2). However, instead of the integrability conditions on ψ, f_t, g_t, we now assume the following.

Assumption 4.1.4. The initial data ψ is an $\mathcal{F}_0 \times \mathcal{B}(\mathbb{R}^d)$-measurable mapping from $\Omega \times \mathbb{R}^d$ to $\mathbb{R}$, such that $\psi \in P_s^m$ (a.s.). The free data f and g are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable mappings from $\Omega \times [0,T] \times \mathbb{R}^d$ to $\mathbb{R}$ and l_2, respectively. Moreover, almost surely (f_t) is a P_s^m-valued process and (g_t) is a $P_s^m(l_2)$-valued process, such that almost surely

$$\big| \|f_t\|_{P_s^m} + \|g_t\|_{P_s^m(l_2)} \big|_{L_\infty[0,T]} < \infty.$$

Definition 4.1.1. A $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable mapping u from $\Omega \times [0,T] \times \mathbb{R}^d$ to $\mathbb{R}$ such that $(u_t)_{t\in[0,T]}$ is almost surely a P_s^1-valued bounded process, is called a classical solution of (4.1.12)-(4.1.13) on $[0,T]$, if almost surely u and its first and second order partial derivatives in x are continuous functions of $(t,x) \in [0,T]\times\mathbb{R}^d$, and almost surely

$$u_t(x) = \psi(x) + \int_0^t [D_i(a_s^{ij}(x)D_j u_s(x)) + b_s^i(x)D_j u_s(x) + c_s(x)u_s(x) + f_s(x)]\, ds$$

$$+ \int_0^t [\sigma_s^{ir}(x)D_i u_s(x) + \nu_s^r(x)u_s(x) + g_s^r(x)]\, dw_s^r$$

for all $(t, x) \in [0, T] \times \mathbb{R}^d$ for a suitable modification of the stochastic integral in the right-hand side of the equation.

If $m \geq 1$, then as noted above the initial condition and free terms are continuous in space. This makes it reasonable to consider the finite difference scheme (4.1.3)-(4.1.4) as an approximation for the problem (4.1.12)-(4.1.13).

Theorem 4.1.4. *Let $k \geq 0$ be integer, and let $\bar{s} > s \geq 0$ be real numbers. Suppose that Assumptions 4.1.1, 4.1.2, and 4.1.4 hold with $m > 2k + 3$.*

(i) Equation (4.1.12)-(4.1.13) *admits a unique $P_{\bar{s}}^{m-1}$-valued classical solution $(u_t)_{t\in[0,T]}$.*

(ii) For fixed h the corresponding finite difference equation (4.1.3)-(4.1.4) *admits a unique $P_{h,\bar{s}}$-valued solution $(u_t^h)_{t\in[0,T]}$.*

(iii) Suppose furthermore $\mathfrak{p}_h^\gamma \geq \kappa$ for $\gamma \in \Lambda_1$, for some constant $\kappa > 0$, and

$$\Lambda_0 \cup -\Lambda_0 \subset \Lambda_1.$$

Then there are continuous random fields $u^{(1)}, \dots u^{(k)}$ on $[0, T] \times \mathbb{R}^d$, independent of h, such that almost surely

$$u_t^h = \sum_{j=0}^{k} \frac{h^j}{j!} u_t^{(j)}(x) + h^{k+1} r_t^h(x)$$

for $t \in [0, T]$ and $x \in \mathbb{G}_h$, where $u^{(0)} = u$, r^h is a continuous random field on $[0, T] \times \mathbb{R}^d$, and for any $q > 0$

$$E \sup_{t\in[0,T]} \sup_{x\in\mathbb{G}_h} |r_t^h(x)\rho_{\bar{s}}(x)|^q + E \sup_{t\in[0,T]} |r_t^h \rho_{\bar{s}}|_{l_{p,h}}^q$$

$$\leq N \left(E\|\psi\|_{P_s^m}^q + E\big| \|f_t\|_{P_s^m} + \|g_t\|_{P_s^m(l_2)} \big|_{L_\infty[0,T]}^q \right)$$

with some $N = N(K, T, m, s, \bar{s}, q, d, |\Lambda|, \kappa)$.

(iv) *Let $(h_n)_{n=1}^{\infty} \in l_q$ be a nonnegative sequence for some $q \geq 1$. Then for every $\varepsilon, M > 0$ there exists a random variable $\xi_{\varepsilon,M}$ such that almost surely*

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{G}_h,|x|\leq M} |u_t(x) - v_t^h(x)| \leq \xi_{\varepsilon,M} h^{k+1-\varepsilon}$$

for $h = h_n$.

Remark 4.1.6. Condition $\mathfrak{p}_h^{\gamma} \geq \kappa$ in assertion (iii) of the above theorem is harmless, similarly to the second part of (4.1.9). As seen in Example 4.1.1, we can always satisfy this additional requirement by adding a sufficiently large constant to $\mathfrak{p}_h^{\gamma}$.

4.2 L_p estimates and acceleration - Proofs

First let us collect some properties of the finite difference operators. Throughout this section we consider a fixed $h > 0$ and use the notation $u_\alpha = D^\alpha u$. It is easy to see that, analogously to the integration by parts,

$$\int_{\mathbb{R}^d} v(\delta_{h,\lambda} u)\,dx = \int_{\mathbb{R}^d} (\delta_{h,-\lambda} v)u\,dx = -\int_{\mathbb{R}^d} (\delta_{-h,\lambda} v)u\,dx, \tag{4.2.14}$$

when $v \in L_{q/q-1}$ and $u \in L_q$ for some $1 \leq q \leq \infty$, with the convention $1/0 = \infty$ and $\infty/(\infty - 1) = 1$. The discrete analogue of the Leibniz rule can be written as

$$\delta_{h,\lambda}(uv) = u(\delta_{h,\lambda} v) + (\delta_{h,\lambda} u)(T_{h,\lambda} v). \tag{4.2.15}$$

Finally, we will also make use of the simple identities

$$T_{h,\alpha}\delta_{h,\beta} u = \delta_{h,\alpha+\beta} u - \delta_{h,\alpha} u, \tag{4.2.16}$$

$$vv_\lambda = (1/2)(\delta_\lambda(v^2) - h(\delta_\lambda v)^2), \tag{4.2.17}$$

and the estimate

$$|\delta_{h,\lambda} v|_{L_p} \leq |\int_0^1 \partial_\lambda v(\cdot + \theta h\lambda)\,d\theta|_{L_p} \leq |\lambda||v|_{W_p^1}, \tag{4.2.18}$$

valid for $p \in [1, \infty]$ and $v \in W_p^1$, $h \neq 0$ and $\lambda \in \mathbb{R}^d$.

Lemma 4.2.1. *For any $p = 2k$ with an integer k, $\lambda \in \mathbb{R}^d$, $h \neq 0$ and real function v on $\mathbb{R}^d$ we can write*

$$\delta_{h,\lambda}(v^{p-1}) = F_p^{h,\lambda}(v)\delta_{h,\lambda} v,$$

where $F_p^{h,\lambda}(v) \geq (1/2)v^{p-2}$ *on* $\mathbb{R}$. *Moreover, for* $p > 2$, $q = p/(p-2)$ *and for all* $v \in L_p(\mathbb{R}^d)$

$$|F_p^{h,\lambda}(v)|_{L_q}^q \leq (p-1)|v|_{L_p}^p. \tag{4.2.19}$$

Proof. First we claim that

$$F_p^{h,\lambda}(v) = \sum_{i=0}^{p-2} v^{p-2-i} T_{h,\lambda} v^i. \tag{4.2.20}$$

This is trivial for $p = 2$, and we have, using (4.2.15)

$$\delta_{h,\lambda}(v^{p-1}) = \delta_{h,\lambda}(v^2 v^{p-3}) = v^2 F_{p-2}^{h,\lambda}(v)\delta_{h,\lambda}v + (v\delta_{h,\lambda}v + \delta_{h,\lambda}vT_{h,\lambda}v)T_{h,\lambda}v^{p-3}.$$

Thus by induction we get (4.2.20), and (4.2.19) follows. For the other claim, clearly we have $F_p^{h,\lambda}(v) \geq (1/2)v^{p-2}$ for $p = 2$. Then we can prove by induction once again, as from (4.2.20) we have,

$$F_p^{h,\lambda}(v) = T_{h,\lambda}v^2 F_{p-2}^{h,\lambda}(v) + T_{h,\lambda}vv^{p-3} + v^{p-2}$$

$$\geq T_{h,\lambda}v^2(1/2)v^{p-4} + T_{h,\lambda}vv^{p-3} + v^{p-2} = (1/2)v^{p-4}(T_{h,\lambda}v + v)^2 + (1/2)v^{p-2}.$$

□

Introduce the notation

$$\mathfrak{A}_q z = -\mathfrak{a}_h^\lambda (z_\lambda)^2 + q\mathfrak{s}_h^{\lambda r}\mathfrak{s}_h^{\mu r} z_\lambda z_\mu$$

for $q \geq 0$ and $z = (z_\lambda)_{\lambda \in \Lambda_0}$, and recall that under condition (4.1.9), $\mathfrak{A}_q z \leq 0$ for $q \leq 2p$.

Lemma 4.2.2. *Let* $m \geq 1$ *be an integer and* $p = 2^k$ *for some integer* $k \geq 1$, *and let Assumptions 4.1.2 and 4.1.3, along with the condition* (4.1.9) *with* p *in place of* $2p$ *be satisfied. Then for* $u \in W_p^m$, $f \in W_p^m$, $g \in W_p^{m+1}(l_2)$ *and for all multi-indices* α *of length* $|\alpha| \leq m$ *we have*

$$\int_{\mathbb{R}^d} p u_\alpha^{p-1}(x) D^\alpha(L_t^h u(x) + f(x))$$

$$+(1/2)p(p-1)u_\alpha^{p-2}(x)|D^\alpha(M_t^{hr}u(x) + g^r(x))|^2 dx$$

$$\leq N \int_{\mathbb{R}^d} u_\alpha^{p-2}(x)\mathfrak{A}_{p-1/4}(\delta_{h,\lambda}u_\alpha(x))_{\lambda\in\Lambda_0}\, dx + N(|u|_{W_p^m}^p + |f|_{W_p^m}^p + |g|_{W_p^m(l_2)}^p) \tag{4.2.21}$$

for $P\times dt$-almost all $(\omega,t)\in\Omega\times[0,T]$, where N is a constant depending only on $d,p,m,|\Lambda|$, and K.

Proof. For real functions v and w defined on $\mathbb{R}^d$ we write $v\sim w$ if their integrals over $\mathbb{R}^d$ are the same. We use the notation $v\preceq w$ if $v\le w+F$ with a function F whose integral over $\mathbb{R}^d$ can be estimated by $N(|u|^p_{W^m_p}+|f|^p_{W^m_p}+|g|^p_{W^m_p(l_2)})$.

By Hölder's inequality we get

$$u_\alpha^{p-1}D^\alpha(\mathfrak{c}_h^\gamma T_{h,\gamma}u)+u_\alpha^{p-1}f_\alpha+u_\alpha^{p-2}|D^\alpha(\mathfrak{n}_h^{\gamma r}T_{h,\gamma}u+g^r)|^2\preceq 0.$$

Next, notice that

$$u_\alpha^{p-1}D^\alpha(\mathfrak{p}_h^\lambda\delta_{h,\lambda}u)\preceq u_\alpha^{p-1}\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha.$$

Then we can repeatedly use (4.2.17) and the nonnegativity of $\mathfrak{p}_h^\lambda$ to get

$$u_\alpha^{p-1}\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha\le(1/2)u_\alpha^{p-2}\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha^2$$

$$\le(1/4)u_\alpha^{p-4}\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha^4\le\cdots\le(1/2^k)\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha^{2^k}.$$

By (4.2.14), $\mathfrak{p}_h^\lambda\delta_{h,\lambda}u_\alpha^p\sim\delta_{h,-\lambda}\mathfrak{p}_h^\lambda u_\alpha^p$. Therefore, as by (4.2.18), $|\delta_{h,-\lambda}\mathfrak{p}_h^\lambda|\le K|\lambda|$, we obtain

$$u_\alpha^{p-1}D^\alpha(\mathfrak{p}_h^\lambda\delta_{h,\lambda}u)\preceq 0.$$

The remaining terms will be treated together. First notice that by Young's and Hölders inequalities

$$(1/2)p(p-1)u_\alpha^{p-2}(\sum_{\lambda\in\Lambda_0}D^\alpha\mathfrak{s}_h^{\lambda r}\delta_{h,\lambda}u,\sum_{\lambda\in\Lambda_0}D^\alpha\mathfrak{s}_h^{\lambda r}\delta_{h,\lambda}u+\sum_{\gamma\in\Lambda_1}\mathfrak{n}_h^{\gamma r}T_{h,\gamma}u+g^r)_{l_2}$$

$$\preceq(1+\varepsilon)(1/2)p(p-1)u_\alpha^{p-2}\mathfrak{s}_h^{\lambda r}\mathfrak{s}_h^{\mu r}\delta_{h,\lambda}u_\alpha\delta_{h,\mu}u_\alpha \tag{4.2.22}$$

for any $\varepsilon>0$, in particular, we can make the prefactor less than $(1/2)p(p-1/2)$. Now for a moment assume $m=0$. By (4.2.14) and Lemma 4.2.1

$$u^{p-1}\delta_{-h\lambda}(\mathfrak{a}_h^\lambda\delta_{h,\lambda}u)\sim-\delta_{h,\lambda}(u^{p-1})\mathfrak{a}_h^\lambda\delta_{h,\lambda}u$$

$$=-F_p^{h,\lambda}(u)\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u)^2\le(1/2)u^{p-2}\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u)^2, \tag{4.2.23}$$

where F is the functional obtained from Lemma 4.2.1. Combining this with (4.2.22), the claim follows for $m=0$. Assume now $m\ge1$. Then it is easy to see that

$$u_\alpha^{p-1}D^\alpha\delta_{-h,\lambda}(\mathfrak{a}_h^\lambda\delta_{h,\lambda}u)\preceq I_1+I_2, \tag{4.2.24}$$

with

$$I_1 := u_\alpha^{p-1} \sum_{(\alpha',\alpha'')\in A} \delta_{-h,\lambda} D^{\alpha'} \mathfrak{a}_h^\lambda D^{\alpha''} \delta_{h,\lambda} u$$

$$I_2 := u_\alpha^{p-1} \delta_{-h,\lambda}(\mathfrak{a}_h^\lambda \delta_{h,\lambda} u_\alpha),$$

where A is the set of ordered pairs of multi-indices (α', α'') such that $|\alpha'| = 1$ and $\alpha' + \alpha'' = \alpha$. By (4.2.14) and Lemma 4.2.1

$$I_1 \sim -2F_p^{h,\lambda}(u_\alpha)\sqrt{\mathfrak{a}_h^\lambda}\delta_{h,\lambda}u_\alpha \sum_{(\alpha',\alpha'')\in A} D^{\alpha'}\sqrt{\mathfrak{a}_h^\lambda}\delta_{h,\lambda}u_{\alpha''}$$

$$\leq \varepsilon F_p^{h,\lambda}(u_\alpha)\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u_\alpha)^2 + \varepsilon^{-1} N F_p^{h,\lambda}(u_\alpha)(\delta_{h,\lambda}u_{\alpha''})^2 \tag{4.2.25}$$

for every $\varepsilon > 0$. Using (4.2.23) with u_α in place of u we get

$$I_2 \preceq -F_p^{h,\lambda}(u_\alpha)\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u_\alpha)^2.$$

Combining this with (4.2.25), from (4.2.24) we obtain

$$I \leq -(1-\varepsilon)F_p^{h,\lambda}(u_\alpha)\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u_\alpha)^2 + \varepsilon^{-1}NF_p^{h,\lambda}(u_\alpha)\sum_{(\alpha',\alpha'')\in A}(\delta_\lambda^h u_{\alpha''})^2$$

$$\leq -(1-\varepsilon)(1/2)u_\alpha^{p-2}\mathfrak{a}_h^\lambda(\delta_{h,\lambda}u_\alpha)^2 + \varepsilon^{-1}N\left[|F_p^{h,\lambda}(u_\alpha)|^q + |\sum_{(\alpha',\alpha'')\in A}(\delta_\lambda^h u_{\alpha''})^2|^{p/2}\right],$$

with $q = p/(p-2)$. The quantity in the brackets is $\preceq 0$, due to the estimates (4.2.19) and (4.2.18). Fixing ε so that $1-\varepsilon > (p-1/2)/(p-1/4)$ and combining the above with (4.2.22), the proof is finished. □

Now we are ready to prove the main a priori estimate. To obtain bounds in Sobolev norms we consider (4.1.3)-(4.1.4) as an SDE on the space W_p^m. Clearly, under Assumption 4.1.2 $u \to L_t^h u$ and $u \to M_t^{h,r} u$ are bounded linear operators from W_p^m to W_p^m and to $W_p^m(l_2)$, respectively, with operator norm uniformly bounded in (t, ω). Therefore if Assumption 4.1.3 is also satisfied, (4.1.3)-(4.1.4) is a SDE in the space W_p^m with Lipschitz continuous coefficients. As such, it admits a unique continuous solution.

Theorem 4.2.3. *Let Assumptions 4.1.2 and 4.1.3 hold with $m \geq 1$, and let condition* (4.1.9) *be satisfied. Then* (4.1.3)-(4.1.4) *has a unique continuous W_p^m-valued solution $(u_t^h)_{t\in[0,T]}$, and for each $q > 0$ there exists a constant $N =$*

$N(d, q, p, m, K, T, |\Lambda|)$ such that

$$E \sup_{t \leq T} |u_t^h|^q_{W^m_p} \leq N(E|\psi|^q_{W^m_p} + EF^q_{m,p}(T) + EG^q_{m,p}(T)) \tag{4.2.26}$$

for all $h > 0$.

Proof. By the preceding argument, we need only prove estimate (4.2.26). Fix $m \geq 1$ and first let $p = 2^k$ for some integer $k \geq 1$, and only assume (4.1.9) with p in place of $2p$, along with Assumptions 4.1.2 and 4.1.3. Let α be a multi-index such that $|\alpha| \leq m$. If we apply Itô's formula to $|D^\alpha u^h|^p_{L_p}$ by Lemma 5.1 in [20], one can notice that the term appearing in the drift is the left-hand side of (4.2.21), with u^h in place of u. Using Corollary 4.2.2 and summing over $|\alpha| \leq m$ we get

$$\begin{aligned} d|u_t^h|^p_{W^m_p} \leq N \int_{\mathbb{R}^d} (D^\alpha u_t^h)^{p-2} \mathfrak{A}_{p-1/4}(\delta_{h,\lambda} u^h_{t,\alpha})_{\lambda \in \Lambda_0} \, dx \, dt \\ + N(|u_t^h|^p_{W^m_p} + |f_t|^p_{W^m_p} + |g_t|^p_{W^{m+1}_p}) \, dt + dm_t^h \end{aligned} \tag{4.2.27}$$

with some N depending only on p, m, d, $|\Lambda|$, and K, where

$$dm_t^h = (p-1) \int_{\mathbb{R}^d} (D^\alpha u_t^h)^{p-1} D^\alpha (\mathfrak{s}_h^{\lambda r} \delta_{h,\lambda} u_t^h + \mathfrak{n}_h^{\gamma r} T_{h,\gamma} u_t^h + g_t^r) \, dx \, dw_t^r$$

with α also used as a repeated index. It is clear that

$$d\langle m^h \rangle_t = (p-1)^2 \sum_{r=1}^\infty \left(\int_{\mathbb{R}^d} (D^\alpha u_t^h)^{p-1} D^\alpha (\mathfrak{s}_h^{\lambda r} \delta_{h,\lambda} u_t^h + \mathfrak{n}_h^{\gamma r} T_{h,\gamma} u_t^h + g_t^r) dx \right)^2 dt.$$

For $p = 2$ Gronwall's lemma can be readily applied as follows. Since for $v \in W^m_2$, $|\alpha| \leq m$, and any function s with derivatives up to order $m+1$ bounded by K, we have

$$\Big| \int_{\mathbb{R}^d} v_\alpha D^\alpha (s \delta_{h,\lambda} v) \, dx \Big| \leq N |v|^2_{W^m_2},$$

(see [13]), we find that the conditions of Lemma 1.2.5 are satisfied with $y_t = |u_t^h|^2_{W^m_2}$, $F = G = N(|f_t|^2_{W^m_2} + |g_t|^2_{W^m_2})$, and $\rho = 1/2$, and therefore the claim follows for $p = 2$ and arbitrary $q > 0$.

For $p = 2k$, from (4.2.27) from (the classical) Gronwall's lemma we have

$$\begin{aligned} \sup_{t \in [0,T]} E|u_t^h|^p_{W^m_p} + E \int_0^T \int_{\mathbb{R}^d} -(u^h_{t,\alpha})^{p-2} \mathfrak{A}_{p-1/4}(\delta_{h,\lambda} u^h_{t,\alpha})_{\lambda \in \Lambda_0} \, dx \, dt \\ \leq NE(|\psi|^p_{W^m_p} + F^p_{m,p}(T) + G^p_{m,p}(T)). \end{aligned} \tag{4.2.28}$$

Therefore, after taking supremum in (4.2.27) and using the Burkholder-Gundy-

Davis inequality we obtain

$$E\sup_{t\in[0,T]}|u^h_t|^p_{W^m_p}\leq NE(|\psi|^p_{W^m_p}+F^p_{m,p}(T)+G^p_{m,p}(T))+NE\langle m^h\rangle^{1/2}_T$$

By Minkowski's and Young's inequality we have

$$\langle m^h\rangle_T\leq\varepsilon\sup_{t\in[0,T]}|u^h_t|^p_{W^m_p}+\varepsilon^{-1}N\int_0^T\int_{\mathbb{R}^d}(D^\alpha u^h_t)^{p-2}\mathfrak{s}^{\lambda r}_h\delta_{h,\lambda}u^h_{t,\alpha}\mathfrak{s}^{\mu r}_h\delta_{h,\mu}u^h_{t,\alpha}\,dx\,dt$$

$$+\varepsilon^{-1}N\int_0^T|u^h_t|^p_{W^m_p}+|g_t|^p_{W^m_p(l_2)}\,dt. \tag{4.2.29}$$

Noticing that

$$(1/4)\mathfrak{s}^{\lambda r}_h u^h_{t,\alpha}\mathfrak{s}^{\mu r}_h z^\lambda z^\mu\leq(1/4)\mathfrak{s}^{\lambda r}_h u^h_{t,\alpha}\mathfrak{s}^{\mu r}_h z^\lambda z^\mu-\mathfrak{A}_p z\leq-\mathfrak{A}_{p-1/4}z,$$

the expectation of second term on the right-hand side of (4.2.29) can be estimated using (4.2.28). Doing so and choosing ε small enough, we get

$$E\sup_{t\in[0,T]}|u^h_t|^p_{W^m_p}\leq NE(|\psi|^p_{W^m_p}+F^p_{m,p}(T)+G^p_{m,p}(T))+(1/2)E\sup_{t\in[0,T]}|u^h_t|^p_{W^m_p},$$

and since the right hand side is finite, the claim follows, for $p=2^k$, $q=p$.

Note that (4.2.26) is equivalent to

$$[E\sup_{t\leq T}|u^h_t|^q_{W^m_p}]^{\frac{1}{q}}\leq N([E|\psi|^q_{W^m_p}]^{\frac{1}{q}}+[EF^q_{m,p}]^{\frac{1}{q}}+[EG^q_{m,p}]^{\frac{1}{q}}),$$

which implies

$$[E\left(\int_0^T|u^h_t|^r_{W^m_p}\right)^{\frac{q}{r}}]^{\frac{1}{q}}\leq N([E|\psi|^q_{W^m_p}]^{\frac{1}{q}}+[EF^q_{m,p}]^{\frac{1}{q}}+[\ G^q_{m,p}]^{\frac{1}{q}}), \tag{4.2.30}$$

for any $r>1$, with another constant N, independent from r. In other words, this means that for the special cases of p and q considered so far the solution operator

$$(\psi,f,g)\to u^h$$

continuously maps $\Psi^m_{p,q}\times\mathcal{F}^m_{p,q}\times\mathcal{G}^m_{p,q}$ to $\mathcal{U}^m_{p,q}$, where

$$\Psi^m_{p,q}=L_q(\Omega,W^m_p),$$

$$\mathcal{F}^m_{p,q}=L_q(\Omega,L_p([0,T],W^m_p)),$$

$$\mathcal{G}^m_{p,q} = L_q(\Omega, L_p([0,T], W^m_p(l_2))),$$

$$\mathcal{U}^m_{p,q} = L_q(\Omega, L_r([0,T], W^m_p)).$$

Let us denote the complex interpolation space between any Banach spaces A_0 and A_1 with parameter θ by $[A_0, A_1]_\theta$. Recall the following interpolation properties (see 1.9.3, 1.18.4, and 2.4.2 from [37])

(i) If a linear operator T is continuous from A_0 to B_0 and from A_1 to B_1, then it is also continuous from $[A_0, A_1]_\theta$ to $[B_0, B_1]_\theta$, moreover, its norm between the interpolated spaces depends only on θ and its norm between the original spaces.

(ii) For a measure space M and $1 < p_0, p_1 < \infty$,

$$[L_{p_0}(M, A_0), L_{p_1}(M, A_1)]_\theta = L_{p_\theta}(M, [A_0, A_1]_\theta),$$

where $1/p_\theta = (1-\theta)/p_0 + \theta/p_1$.

(iii) For $m \in \mathbb{N}$, $1 < p_0, p_1 < \infty$,

$$[W^m_{p_0}, W^m_{p_1}]_\theta = W^m_{p_\theta},$$

where $1/p_\theta = (1-\theta)/p_0 + \theta/p_1$.

Now take any $p > 2$ and take $p_1 = T(p) := 2^k$ for the smallest k such that $2^k > p$. Define $\theta \in [0,1]$ by $1/p = (1-\theta)/2 + \theta/p_1$. Further, take $q \geq p_1$ and define q_0 with $1/q = (1-\theta)/q_0 + \theta/p_1$. Notice that since (4.1.9) is assumed, (4.1.9) also holds with p_1 in place of $2p$. By properties (ii), (iii), we have

$$\Psi^m_{p,q} = [\Psi^m_{2,q_0}, \Psi^m_{p_1,p_1}]_\theta,\ \mathcal{F}^m_{p,q} = [\mathcal{F}^m_{2,q_0}, \mathcal{F}^m_{p_1,p_1}]_\theta,$$

$$\mathcal{G}^m_{p,q} = [\mathcal{G}^m_{2,q_0}, \mathcal{G}^m_{p_1,p_1}]_\theta,\ \mathcal{U}^m_{p,q} = [\mathcal{U}^m_{2,q_0}, \mathcal{U}^m_{p_1,p_1}]_\theta,$$

and since we know the continuity of the solution operator from $\Psi^m_{2,q_0} \times \mathcal{F}^m_{2,q_0} \times \mathcal{G}^m_{2,q_0}$ to $\mathcal{U}^m_{2,q_0}$ and from $\Psi^m_{p_1,p_1} \times \mathcal{F}^m_{p_1,p_1} \times \mathcal{G}^m_{p_1,p_1}$ to $\mathcal{U}^m_{p_1,p_1}$, by (i), the solution operator is also continuous from $\Psi^m_{p,q} \times \mathcal{F}^m_{p,q} \times \mathcal{G}^m_{p,q}$ to $\mathcal{U}^m_{p,q}$ for any $p \geq 2$, $q \geq T(p)$. Moreover, its norm is independent of r. Hence we have (4.2.30), and letting $r \to \infty$ and keeping in mind that u^h is a continuous in W^m_p-valued process, using Fatou's lemma we get (4.2.26). The case $q < T(p)$ can be covered by the usual application of Lemma 1.2.3. $\square$

Proof of Theorems 4.1.1-4.1.3. To prove Theorem 4.1.1, first consider the functions

$$F(h) = \delta_{h,\lambda}\phi(x) = \int_0^1 \partial_\lambda \phi(x + h\theta\lambda)\, d\theta,$$

$$G(h) = \delta_{-h,\lambda}\delta_{h,\lambda}\psi(x) = \int_{-1}^0 \int_0^1 \partial_\lambda\partial_\lambda \psi(x + h\lambda(\theta_1 + \theta_2)) d\theta_1 d\theta_2$$

for fixed $\phi \in W_p^{n+l+2}, \psi \in W_p^{n+l+3}$, $n, l \geq 0$. Applying Taylor's formula at $h = 0$ up to $n + 1$ terms we get that

$$|\delta_{h,\lambda}\phi - \sum_{i=0}^{n} h^i A_i \partial_\lambda^{i+1}\phi|_{W_p^l} \leq N|h|^{n+1}|\phi|_{W_p^{n+l+2}},$$

$$|\delta_{-h,\lambda}\delta_{h,\lambda}\psi - \sum_{i=0}^{n} h^i B_i \partial_\lambda^{i+2}\psi|_{W_p^l} \leq N|h|^{n+1}|\psi|_{W_p^{l+n+3}}$$

with constants $A_i = 1/(i + 1)!$ and

$$B_i = \begin{cases} 0 & \text{if } i \text{ is odd} \\ \frac{2}{(i+2)!} & \text{if } i \text{ is even} \end{cases},$$

where $N = N(|\Lambda|, d, l, n)$ is a constant. Similarly, or in fact equivalently to the first inequality, we have

$$|T_{h,\lambda}\varphi - \sum_{i=0}^{n} \frac{h^i}{i!}\partial_\lambda^i \varphi|_{W_p^l} \leq N|h|^{n+1}|\varphi|_{W_p^{n+l+1}}$$

for $\varphi \in W_p^{n+l+1}$, where ∂_λ^0 denotes the identity operator. Without going into details, it is clear that, due to Assumption 4.1.2, from these expansions one can obtain operators $\mathfrak{L}_t^{(i)}, \mathfrak{M}_t^{(i)k}$ for integers $i \in [0, m]$ such that $\mathfrak{L}_t^0\phi = \partial_i a^{ij}\partial_j\phi + b^i\partial_i\phi + c\phi$, $\mathfrak{M}_t^{(0)k}\phi = \sigma^{ik}\partial_i\phi + \nu^k\phi$,

$$|\mathfrak{L}_t^{(i)}\phi|_{W_p^l} \leq N|\phi|_{W_p^{l+i+1}} \quad \text{for } i \text{ odd, } i + l \leq m, \tag{4.2.31}$$

$$|\mathfrak{L}_t^{(i)}\phi|_{W_p^l} \leq N|\phi|_{W_p^{l+i+2}} \quad \text{for } i \text{ even, } i + l \leq m, \tag{4.2.32}$$

$$|\mathfrak{M}_t^{(i)}\phi|_{W_p^l(l_2)} \leq N|\phi|_{W_p^{l+i+1}} \quad i + l \leq m \tag{4.2.33}$$

and

$$|(L_t^h - \sum_{i=0}^{n} \frac{h^i}{i!}\mathfrak{L}_t^{(i)})\phi|_{W_p^l} \leq N|h|^{n+1}|\phi|_{W_p^{n+l+3}} \quad \text{for } n + l < m, \tag{4.2.34}$$

$$|(M^h_t - \sum_{i=0}^{n} \frac{h^i}{i!}\mathfrak{M}^{(i)}_t)\phi|_{W^l_p(l_2)} \le N|h|^{n+1}|\phi|_{W^{n+l+2}_p} \quad \text{for } n+l<m \tag{4.2.35}$$

with $N = N(|\Lambda|, K, d, p, m)$. The random fields $u^{(j)}$ in expansion (4.1.10) can then be obtained from the system of SPDEs

$$du^{(j)}_t = (\mathfrak{L}^{(0)}_t u^{(j)}_t + \sum_{l=1}^{j} \tbinom{j}{l}\mathfrak{L}^{(l)}_t u^{(j-l)}_t)\,dt + (\mathfrak{M}^{(0)r}_t u^{(j)}_t + \sum_{l=1}^{j} \tbinom{j}{l}\mathfrak{M}^{(l)r}_t u^{(j-l)}_t)\,dw^r_t \tag{4.2.36}$$

$$u^{(j)}_0 = 0, \quad j = 1, ..., k, \tag{4.2.37}$$

where $v^{(0)} = u$, the solution of (4.1.1)-(4.1.2). The following theorem holds, being the exact analogue of Theorem 5.1 from [13]. It can be proven inductively on j, by a straightforward application of Theorem 3.1.1 and (4.2.31)-(4.2.33).

Theorem 4.2.4. *Let $k \ge 1$ be an integer, and let Assumptions 4.1.1, 4.1.2 and 4.1.3 hold with $m \ge 2k+1$. Then there is a unique solution $u^{(1)}, \dots, u^{(k)}$ of (4.2.36)-(3.1.2). Moreover, $u^{(j)}$ is a W^{m-2j}_p-valued weakly continuous process, it is strongly continuous as a W^{m-2j-1}_p-valued process, and*

$$E \sup_{t\in[0,T]} |u^{(j)}_t|^q_{W^{m-2j}_p} \le N(E|\psi|^q_{W^m_p} + EF^q_{m,p}(T) + EG^q_{m,p}(T))$$

for $j = 1, \dots, k$, for any $q > 0$, with a constant $N = N(K, m, p, q, T, |\Lambda|)$.

Set

$$\bar{r}^h_t(x) = u^h_t(x) - \sum_{j=0}^{k} \frac{h^j}{j!} u^{(j)}_t(x),$$

for $t \in [0,T]$ and $x \in \mathbb{R}^d$, where u^h is the W^m_p-valued solution of (4.1.3)-(4.1.4). Then it is not difficult to verify that $\bar{r}^h$ is the solution, of the finite difference equation

$$\bar{r}^h_t(x) = (L^h_t \bar{r}^h_t(x) + F^h_t(x))\,dt + (M^{hr}_t \bar{r}^h_t(x) + G^{hr}_t(x))\,dw^r_t, \quad t \in (0,T],\, x \in \mathbb{R}^d$$

with initial condition $\bar{r}^h_0(x) = 0$ for $x \in \mathbb{R}^d$, where

$$F^h_t = \sum_{j=0}^{k} \frac{h^j}{j!}\left(L^h_t - \sum_{i=0}^{k-j} \frac{h^i}{i!}\mathfrak{L}^{(i)}_t\right) u^{(j)}_t,$$

$$G^h_t = \sum_{j=0}^{k} \frac{h^j}{j!}\left(M^{hr}_t - \sum_{i=0}^{k-j} \frac{h^i}{i!}\mathfrak{M}^{(i)r}_t\right) u^{(j)}_t.$$

Hence by applying Theorem 4.2.3 we get

$$E\sup_{t\in[0,T]}|\overline{r}_t^h|^q_{W_p^{m-2k-3}} \le NE\left(\int_0^t |F_t|^p_{W_p^{m-2k-3}} + |G_t|^p_{W_p^{m-2k-3}}\,dt\right)^{q/p}.$$

Now using $m-2k-3>d/p$, for the left-hand side we can write

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|\overline{r}_t^h(x)|^q + E\sup_{t\in[0,T]}|\overline{r}_t^h|^q_{l_{p,h}} \le NE\sup_{t\in[0,T]}|\overline{r}_t^h|^q_{W_p^{m-2k-3}},$$

while (4.2.34), (4.2.35), and the theorem above yield

$$E\sup_{t\in[0,T]}|F_t^h|^q_{W_p^{m-2k-3}} + |G_t^h|^q_{W_p^{m-2k-3}} \le Nh^{q(k+1)}\sum_{j=0}^k E\sup_{t\in[0,T]}|u_t^{(j)}|^q_{W_p^{m-2j}}$$

$$\le Nh^{q(k+1)}(E|\psi|^q_{W_p^m} + EF^q_{m,p}(T) + EG^q_{m,p}(T)),$$

where N denotes some constants which depend only on K, m, d, q, p, T and $|\Lambda|$. Putting these inequalities together we obtain the estimate

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|\overline{r}_t^h(x)|^q + E\sup_{t\in[0,T]}|\overline{r}_t^h|^q_{l_{p,h}}$$

$$\le Nh^{q(k+1)}(E|\psi|^q_{W_p^m} + EF^q_{m,p}(T) + EG^q_{m,p}(T)), \tag{4.2.38}$$

for all $h>0$ with a constant $N=N(K,m,d,q,p,T,|\Lambda|)$. Thus we have the following theorem.

Theorem 4.2.5. *Let $k\ge 0$ be an integer and let Assumptions 4.1.1, 4.1.2 and 4.1.3 hold with $m>2k+3+d/p$. Then there are continuous random fields $u^{(1)},\dots u^{(k)}$ on $[0,T]\times\mathbb{R}^d$, independent of h, such that almost surely*

$$u_t^h(x) = \sum_{j=0}^k \frac{h^j}{j!}u_t^{(j)}(x) + \overline{r}_t^h(x) \tag{4.2.39}$$

for all $t\in[0,T]$ and $x\in\mathbb{R}^d$, where $u^{(0)}=u$, u^h is the W_p^m-valued solution of (4.1.3)-(4.1.4)*, and $\overline{r}^h$ is a continuous random field on $[0,T]\times\mathbb{R}^d$, which for any $q>0$ satisfies the estimate* (4.2.38)*.*

Proof. The expansion (4.2.39) holds by the definition of $\overline{r}^h$, its continuity is a simple consequence of Sobolev embeddings and Theorems 3.1.1, 4.2.3 and 4.2.4, and estimate (4.2.38) is proved above. □

To finish the proof of Theorem 4.1.1 we need only show that under the conditions of Theorem 4.2.5 the restriction of the W_p^m-valued solution u^h of (4.1.3)-(4.1.4) onto $[0,T] \times \mathbb{G}_h$ is a continuous $l_{p,h}$-valued process which solves (4.1.3)-(4.1.4). To this end note that under the conditions of Theorem 4.2.5 u^h is a continuous W_p^{m-1} valued process, and therefore by (4.1.5) its restriction to $[0,T]\times\mathbb{G}_h$ is a continuous $l_{p,h}$-valued process. To see that this process satisfies (4.1.3)-(4.1.4) we fix a point $x \in \mathbb{G}_h$ and take a nonnegative smooth function φ with compact support in $\mathbb{R}^d$ such that its integral over $\mathbb{R}^d$ is one. Define for each integer $n \geq 1$ the function $\varphi^{(n)}(z) = n^d \varphi(n(z-x))$, $z \in \mathbb{R}^d$. Then we have for u^h, the W_p^m-valued solution of (4.1.3)-(4.1.4), that almost surely

$$(u_t^h, \varphi^{(n)}) = (\psi, \varphi^{(n)}) + \int_0^t (L_s^h u_s^h + f_s, \varphi^{(n)})\, ds + \int_0^t (M_s^{hr} u_s^h + g_s^r, \varphi^{(n)})\, dw_s^r$$

for all $t \in [0,T]$ and for all $n \geq 1$. Letting here $n \to \infty$, for each $t \in [0,T]$ we get

$$u_t^h(x) = \psi(x) + \int_0^t (L_s^h u_s^h(x) + f_s(x))\, ds + \int_0^t (M_s^{hr} u_s^h(x) + g_s^r(x))\, dw_s^r \quad (4.2.40)$$

almost surely, since u^h, ψ, f,, g and the coefficients of L^h and M^h are continuous in x, due to Sobolev's theorem on embedding $W_p^m(\mathbb{R}^d)$ into $C_b(\mathbb{R}^d)$ in the case $m > d/p$. Note that both $u_t^h(x)$ and the random field on the the right-hand side of equation (4.2.38) are continuous in $t \in [0,T]$. Therefore we have this equality almost surely for all $t \in [0,T]$ and $x \in \mathbb{G}_h$. The proof of Theorem 4.1.1 is complete.

The extrapolation result, Theorem 4.1.2, follows from Theorem 4.1.1 by standard calculations, and hence Theorem 4.1.3 on the rate of almost sure convergence follows by a standard application of the Borel-Cantelli Lemma, for further details we refer to [13].

□

Proof of Theorem 4.1.4. Let $\rho(x) = \rho_{\bar{s}}(\epsilon x) = 1/(1+|\epsilon x|^2)^{\bar{s}/2}$, where $\epsilon > 0$ is to be determined later and choose p large enough so that $1 > d/p$ - and therefore $m > 2k+3+d/p$ -, and Assumption 4.1.3 holds for $\psi\rho$, $f\rho$ and $g\rho$ in place of ψ, f and g, respectively. After some calculations one gets that u is the solution of (4.1.12)-(4.1.13) if and only if $u\rho$ is the solution of the equation

$$dv_t(x) = (D_i \hat{a}_t^{ij}(x) D_j v_t(x) + \hat{b}_t^i(x) D_i v_t(x) + \hat{c}_t(x) v_t(x) + f_t \rho(x))\, dt$$

$$+ (\hat{\sigma}_t^{ir}(x) v_t(x) + \hat{\nu}_t^r(x) v_t(x) + g_t^r \rho(x))\, dw_t^r \quad (4.2.41)$$

for $(t,x) \in [0,T] \times \mathbb{R}^d$, with the initial condition

$$v_0(x) = \psi\rho(x), \tag{4.2.42}$$

for $x \in \mathbb{R}^d$, where the coefficients are given by

$$\hat{a}^{ij} = a^{ij},$$

$$\hat{b}^i = b_t^i - 2\sum_{j=1}^d a^{ij}\frac{D_j\rho}{\rho},$$

$$\hat{c} = c - \sum_{i,j=1}^d a^{ij}\frac{D_iD_j\rho}{\rho} - \sum_{i,j=1}^d D_ia^{ij}\frac{D_j\rho}{\rho} - \sum_{i=1}^d \hat{b}^i\frac{D_i\rho}{\rho},$$

$$\hat{\sigma}^{ir} = \sigma^{ir},$$

$$\hat{\nu}^r = \nu^r - \sum_{i=1}^d \sigma^{ir}\frac{D_i\rho}{\rho}.$$

Due to our choice of ρ, these coefficients still satisfy the conditions of Theorem 3.1.1. Applying this theorem, we obtain a W_p^m-valued unique solution v. Using Sobolev embedding, we get that v/ρ - which is a solution of (4.1.12) - is a $P_{\bar{s}}^{m-1}$-valued process.

One can similarly transform the finite difference equations, using (4.2.15)-(4.2.16). It turns out that u^h is a solution of (4.1.3)-(4.1.4) if and only if $u^h\rho$ is a solution of the equation

$$v_t^h(x) = \{\hat{L}_t^h(x)v_t^h(x) + f_t\rho(x))\,dt + (\hat{M}_t^{hr}(x)v_t^h(x) + g_t^r\rho(x))\,dw_t^r \tag{4.2.43}$$

for $(t,x) \in [0,T] \times \mathbb{G}_h$ with initial condition

$$v_0^h(x) = \psi\rho(x), \tag{4.2.44}$$

for $x \in \mathbb{G}_h$, where

$$\hat{L}_t^h\varphi = \sum_{\lambda\in\Lambda_0}\delta_{-h,\lambda}(\hat{\mathfrak{a}}_h^\lambda\delta_{h,\lambda}\varphi) + \sum_{\gamma\in\Lambda_1}\hat{\mathfrak{p}}_h^\gamma\delta_{h,\gamma}\varphi + \sum_{\gamma\in\Lambda_1}\hat{\mathfrak{c}}_h^\gamma T_{h,\gamma}\varphi,$$

$$\hat{M}_t^{hr}\varphi = \sum_{\lambda\in\Lambda_0}\hat{\mathfrak{s}}_h^{\lambda r}\delta_{h,\lambda}\varphi + \sum_{\gamma\in\Lambda_1}\hat{\mathfrak{n}}_h^{\gamma r}T_{h,\gamma}\varphi,$$

with

$$\hat{\mathfrak{a}}_h^\lambda = \mathfrak{a}_h^\lambda,$$

$$\hat{\mathfrak{p}}_h^\lambda = \mathfrak{p}_h^\gamma + \frac{(T_{h,-\lambda}\mathfrak{a}^\lambda)\delta_{h,-\lambda}\rho - (T_{h,\lambda}\mathfrak{a}^{-\lambda})\delta_{h,\lambda}\rho}{\rho},$$

$$\hat{\mathfrak{c}}_h^\lambda = \mathfrak{c}_h^\lambda \frac{\rho}{T_{h,\lambda}\rho} - \frac{(\delta_{h,-\lambda}\mathfrak{a}^\lambda)\delta_{h,-\lambda}\rho - \mathfrak{a}^\lambda\delta_{h,-\lambda}\delta_{h,\lambda}\rho + \hat{\mathfrak{p}}^\lambda\delta_{h,\lambda}\rho}{T_{h,\lambda}\rho},$$

$$\hat{\mathfrak{s}}_h^{\lambda r} = \mathfrak{s}_h^{\lambda r},$$

$$\hat{\mathfrak{n}}_h^{\gamma r} = \mathfrak{n}_h^{\gamma r}\frac{\rho}{T_{h,\lambda}\rho} - \frac{\mathfrak{s}_h^{\lambda r}\delta_{h,\lambda}\rho}{T_{h,\lambda}\rho},$$

where $\mathfrak{a}^\lambda$ is understood to be 0 when not defined.

As it was mentioned earlier, the restriction to $\mathbb{G}_h$ of the continuous modifications of $\psi\rho, f\rho, g\rho$ are in $l_{p,h}$, $l_{p,h}$-valued, and $l_{p,h}(l_2)$-valued processes, respectively. The coefficients above are bounded, so as we have already seen, there exists a unique $l_{p,h}$-valued solution v^h, in particular, it is bounded. Therefore v^h/ρ is a solution of (4.1.3)-(4.1.4) and has polynomial growth.

By choosing ϵ small enough, $|\delta_{h,\lambda}\rho/\rho|$ can be made arbitrarily small, uniformly in $x \in \mathbb{R}^d, \lambda \in \Lambda, |h| < 1$. In particular, we can choose it to be small enough such that $\hat{\mathfrak{p}}_h^\gamma \geq 0$. Moreover, all of the smoothness and boundedness properties of the coefficients are preserved. Therefore (4.2.43)-(4.2.44) is a finite difference scheme for the equation (4.2.41)-(4.2.42) such that it satisfies Assumptions 4.1.1 through 4.1.3. Claims (iii) and (iv) then follow from applying Theorems 4.1.1 and 4.1.3.

4.3 Localization error

Here it will be more convenient to discuss equations in the non-divergence form, that is,

$$du_t(x) = (Lu_t(x) + f_t(x))\,dt + \sum_{k=1}^{\infty}(M^k u_t(x) + g_t^k(x))\,dw_t^k \tag{4.3.45}$$

on $(t,x) \in [0,T] \times \mathbb{R}^d =: H_T$, with initial condition

$$u_0(x) = \psi(x), \quad x \in \mathbb{R}^d, \tag{4.3.46}$$

where

$$L = a_t^{ij}(x)D_iD_j + b_t^i(x)D_i + c_t(x), \quad M^k = \sigma_t^{ik}(x)D_i + \mu_t^k(x), \quad k = 1,2,...,$$

The following assumptions almost coincide with the ones in Chapter 3, we formulate them here for the convenience of the reader, and more importantly, because of the additional assumption on the nonnegative square root ρ of

$$\alpha^{ij} = 2a^{ij} - \sigma^{ik}\sigma^{jk},$$

see Assumption 4.3.2 (c) below. Concerning this assumption we remark that is well-known from [7] that ρ is Lipschitz continuous in x if α is bounded and has bounded second order derivatives, but it is also known that the second order derivatives of ρ may not exist in the classical sense, even if α is smooth with bounded derivatives of arbitrary order.

Assumption 4.3.1. For $P \otimes dt \otimes dx$-almost all $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$

$$\alpha_t^{ij}(x)z^i z^j \geq 0$$

for all $z \in \mathbb{R}^d$.

Assumption 4.3.2. (a) The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m, 2)$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i, j \in \{1, 2, ..., d\}$.

(b) The derivatives in $x \in \mathbb{R}^d$ of b^i and c up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i \in \{1, 2, ..., d\}$. The functions $\sigma^i = (\sigma^{ik})_{k=1}^{\infty}$ and $\mu = (\mu^k)_{k=1}^{\infty}$ are l_2-valued and their derivatives in x up to order $m+1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued functions, bounded by K.

(c) The derivatives in $x \in \mathbb{R}^d$ of $\rho = \sqrt{\alpha}$ up to order $m + 1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K.

Assumption 4.3.3. The initial value, ψ is an $\mathcal{F}_0$-measurable random variable with values in W_p^m. The free data, f_t and $g_t = (g^k)_{k=1}^{\infty}$ are predictable processes with values in W_p^m and $W_p^{m+1}(l_2)$, respectively, such that almost surely

$$\mathcal{K}_{m,p}^p(T) := |\psi|_{W_p^m}^p + \int_0^T \left(|f_t|_{W_p^m}^p + |g_t|_{W_p^{m+1}}^p\right) dt < \infty. \tag{4.3.47}$$

Let us refer to the problem (4.3.45)-(4.3.46) as Eq($\mathfrak{D}$), where $\mathfrak{D}$ stands for the "data"

$$\mathfrak{D} = (\psi, a, b, c, \sigma, \mu, f, g)$$

with $a = (a^{ij})$, $b = (b^i)$, $\sigma = (\sigma^{ki})$, $g = (g^k)$ and $\mu = (\mu^k)$. We are interested in the error when instead of Eq($\mathfrak{D}$) we solve Eq($\bar{\mathfrak{D}}$) with

$$\bar{\mathfrak{D}} = (\bar{\psi}, \bar{a}, \bar{b}, \bar{c}, \bar{\sigma}, \bar{\mu}, \bar{f}, \bar{g}).$$

Assumption 4.3.4. Almost surely

$$\mathfrak{D} = \bar{\mathfrak{D}} \quad \text{on } [0,T] \times \{x \in \mathbb{R}^d : |x| \le R\}. \tag{4.3.48}$$

The main example to keep in mind is when each component of $\bar{\mathfrak{D}}$ is a truncation of the corresponding component of $\mathfrak{D}$. Let $B_R = \{x \in \mathbb{R}^d : |x| \le R\}$ for $R > 0$. Define $\bar{\mathcal{K}}^p_{m,p}(T)$ as $\mathcal{K}^p_{m,p}(T)$ with $\bar{\psi}$, $\bar{f}$ and $\bar{g}$ in place of ψ, f and g, respectively. The main result reads as follows.

Theorem 4.3.1. *Let $\nu \in (0,1)$ and let Assumptions 4.3.1, 4.3.2 (b)-(c) and 4.3.3 hold with $m > 2 + d/p$ for $\mathfrak{D}$ and $\bar{\mathfrak{D}}$. Let also Assumption 4.3.4 hold. Then Eq($\mathfrak{D}$) and Eq($\bar{\mathfrak{D}}$) have a unique classical solution u and $\bar{u}$, respectively, and for $q > 0$, $r' > 1$*

$$E \sup_{t\in[0,T]} \sup_{x\in B_{\nu R}} |u_t(x) - \bar{u}_t(x)|^q \le N e^{-\delta R^2} E^{1/r'}(\mathcal{K}^{qr'}_{m,p}(T) + \bar{\mathcal{K}}^{qr'}_{m,p}(T)), \tag{4.3.49}$$

where N and δ are positive constants, depending on K, d, T, q, r', m, p, and ν.

First we collect some auxiliary results. The following lemma is a version of Kolmogorov's continuity criterion, see Theorem 3.4. of [8].

Lemma 4.3.2. *Let $x(\theta)$ be a stochastic process parametrized by and continuous in $\theta \in D \subset \mathbb{R}^p$, where D is a direct product of lower dimensional closed balls. Then for all $0 < \alpha < 1$, $q \ge 1$, and $s > p/\alpha$,*

$$\mathbb{E} \sup_{\theta} |x(\theta)|^q \le N(1 + |D|) \left[\sup_{\theta} (\mathbb{E}|x(\theta)|^{qs})^{1/s} + \sup_{\theta\neq\theta'} \left(\frac{\mathbb{E}|x(\theta) - x(\theta')|^{qs}}{|\theta - \theta'|^{qs\alpha}} \right)^{1/s} \right]$$

where $N = N(q, s, \alpha, p)$, and $|D|$ is the volume of D.

Lemma 4.3.3. *Let $(\alpha_t)_{t\in[0,T]}$ and $(\beta_t)_{t\in[0,T]}$ be $\mathcal{F}_t$-adapted processes with values in $\mathbb{R}^d$ and $l_2(\mathbb{R}^d)$, respectively, in magnitude bounded by a constant K. Then for the process*

$$X_t = \int_0^t \alpha_s \, ds + \int_0^t \beta^k_s \, dw^k_s, \quad t \in [0,T] \tag{4.3.50}$$

there exist constants $\varepsilon = \varepsilon(K,T) > 0$ and a $N = N(K,T)$ such that

$$E \sup_{t\le T} e^{\varepsilon |X_t|^2} \le N.$$

Proof. By Itô's formula

$$Y_t := e^{|X_t|^2 e^{-\mu t}} = 1 + \int_0^t e^{|X_s|^2 e^{-\mu s} - \mu s} \{|\beta_s|^2 + 2\alpha_s X_s$$

$$+2|\beta_s X_s|^2 - \mu|X_s|^2\}\,ds + m_t$$

for any $\mu \in \mathbb{R}$, where $(m_t)_{t\in[0,T]}$ is a local martingale starting from 0. By simple inequalities

$$2\alpha X + 2|\beta X|^2 \le |\alpha|^2 + |X|^2 + 2|\beta|^2|X|^2 \le K^2 + (2K^2+1)|X|^2.$$

Hence for $\mu = (2K^2+1)$ and for a stopping time $\tau \le T$ we have

$$EY_{t\wedge\tau_n} \le 1 + 2K^2 \int_0^t EY_{s\wedge\tau_n}\,ds,$$

for $\tau_n = \tau \wedge \rho_n$, where $(\rho_n)_{n=1}^\infty$ is a localising sequence of stopping times for m. Hence, by Gronwall's lemma,

$$EY_{t\wedge\tau_n} \le e^{2K^2T}\,.$$

where N is independent from n. Letting here $n \to \infty$, by Fatou's lemma we get

$$Ee^{|X_\tau|^2 e^{-\mu T}} \le Ee^{|X_\tau|^2 e^{-\mu\tau}} \le e^{K^2T}$$

for stopping times $\tau \le T$. Hence applying Lemma 1.2.3 for $r \in (0,1)$ we obtain

$$E\sup_{t\le T} e^{r|X_\tau|^2 e^{-\mu T}} \le Ne^{rK^2T}\,.$$

□

To formulate our next lemma we consider the stochastic differential equation

$$dX_s = \alpha_s(X_s)\,ds + \beta_s^k(X_s)\,dw_s^k, \tag{4.3.51}$$

where α and $\beta = (\beta^k)$ are $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable function on $\Omega \times [0,T] \times \mathbb{R}^d$, with values in $\mathbb{R}^d$ and $l_2(\mathbb{R}^d)$ such that they are bounded in magnitude by K and satisfy the Lipschitz condition in $x \in \mathbb{R}^d$ with a Lipschitz constant M, uniformly in the other arguments. Then equation (4.3.51) with initial condition $X_t = x$ has a unique solution $X^{t,x} = (X_s^{t,x})_{s\in[t,T]}$ for any $t \in [0,T]$ and $x \in \mathbb{R}^d$.

Remark 4.3.1. It is well known from [28] that the solution of (4.3.50) can be chosen to be continuous in t, x, s. In the following, by $X_s^{t,x}$ we always understand such a continuous modification.

Lemma 4.3.4. *Set $\hat{X}^{t,x} = X^{t,x} - x$. Then for any R,*

$$E \sup_{0\le t\le s\le T} \sup_{|x|\le R} e^{|\hat{X}^{t,x}_s|^2\delta} \le N(1+R^{d+1/2}), \tag{4.3.52}$$

and for any R and r

$$P(\sup_{0\le t\le s\le T} \sup_{|x|\le R} |\hat{X}^{t,x}_s| > r) \le N e^{-\delta r^2}(1+R^{d+1/2}), \tag{4.3.53}$$

where $\delta = \delta(d,K,M,T) > 0$ and $N = N(d,K,M,T)$.

Proof. It is easy to see that (4.3.52) implies (4.3.53), so we need only prove the former. For a fixed δ, to be chosen later, let us use the notations $f(y) = e^{|y|^2\delta}$ and $\gamma = 2(d+2)+1$. By Lemma 4.3.2, we have

$$\mathbb{E} \sup_{0\le t\le s\le T} \sup_{|x|\le R} f(\hat{X}^{t,x}_s) \le N(1+R^d) \sup_{0\le t\le s\le T} \sup_{|x|\le R} (\mathbb{E} f^\gamma(\hat{X}^{t,x}_s))^{1/\gamma}$$

$$+ N(1+R^d) \sup_{\substack{0\le t\le s\le T\\ 0\le t'\le s'\le T}} \sup_{\substack{|x|\le R\\ |x'|\le R}} \left(\frac{\mathbb{E}|f(\hat{X}^{t,x}_s) - f(\hat{X}^{t',x'}_{s'})|^\gamma}{(|t-t'|^2 + |s-s'|^2 + |x-x'|^2)^{\gamma/4}} \right)^{1/\gamma}. \tag{4.3.54}$$

The first term above, by Lemma 4.3.3, provided $\delta \le \varepsilon/\gamma$, can be estimated by NR^d. As for the second one,

$$f(\hat{X}^{t,x}_s) - f(\hat{X}^{t',x'}_{s'}) = \int_0^1 \partial f(\vartheta \hat{X}^{t,x}_s + (1-\vartheta)\hat{X}^{t',x'}_s)(\hat{X}^{t,x}_s - \hat{X}^{t',x'}_{s'})\, d\vartheta.$$

Notice that $|\nabla f(y)| \le N(\delta) f^2(y)$, therefore by Jensen's inequality and Lemma 4.3.3 again, provided $\delta \le \varepsilon/(8\gamma)$, we obtain

$$\mathbb{E}|f(\hat{X}^{t,x}_s) - f(\hat{X}^{t',x'}_{s'})|^\gamma \le N \mathbb{E}^{1/2} |\hat{X}^{t,x}_s - \hat{X}^{t',x'}_{s'}|^{2\gamma}.$$

Now the the right-hand side can be estimated by standard moment bounds for SDEs, see e.g. Corollary 2.5.5 in [22], from which we obtain

$$\left(\frac{\mathbb{E}|f(\hat{X}^{t,x}_s) - f(\hat{X}^{t',x'}_{s'})|^{2\gamma}}{(|t-t'|^2 + |s-s'|^2 + |x-x'|^2)^{\gamma/2}} \right)^{1/(2\gamma)} \le N(1+R^{1/2}).$$

□

Proof of Theorem 4.3.1. Throughout the proof we will use the constant $\lambda = \lambda(d,q)$, which stands for a power of R, and, like N and δ, may change from line to line. Clearly it suffices to prove Theorem 4.3.1 with $e^{-\delta R^2}R^\lambda$ in place of $e^{-\delta R^2}$ in the right-hand side of inequality (4.3.49). We also assume first that $q > 10$.

The main idea of the proof is based on stochastic representation of solutions to linear stochastic PDEs of parabolic type, see [28], [27].

Recall that $\rho = (\rho_t^{ir}(x))_{i,r=1}^d$ is the symmetric nonnegative square root of $\alpha = (2a^{ij} - \sigma^{ik}\sigma^{jk})_{i,j=1}^d$ and $\bar\rho$ is the symmetric nonnegative square root of $\bar\alpha = (2\bar a^{ij} - \bar\sigma^{ik}\bar\sigma^{jk})_{i,j=1}^d$. Then due to Assumption 4.3.4, $\rho = \bar\rho$ almost surely for all $t \in [0, T]$ and for $|x| \le R$.

Let $(\hat w_t^r)_{t\ge 0, r=1\ldots d}$ be a d-dimensional Wiener process, also independent of the σ-algebra $\mathcal{F}_\infty$ generated by $\mathcal{F}_t$ for $t \ge 0$. Consider the problem

$$\begin{aligned} dv_t(x) =&(Lv_t(x) + f_t(x))\,dt + (M^k v_t(x) + g_t^k(x))\,dw_t^k \\ &+ \mathcal{N}^r v_t(x)\,d\hat w_t^r && (4.3.55) \\ v_0(x) =&\psi(x), && (4.3.56) \end{aligned}$$

where $\mathcal{N}^r = \rho^{ri}D_i$. Then by Theorem 3.1.1 and by Sobolev embeddings, (4.3.55)-(4.3.56) has a unique classical solution v, and for each $t \in [0, T]$ and $x \in \mathbb{R}^d$ almost surely

$$u_t(x) = E(v_t(x)|\mathcal{F}_t). \tag{4.3.57}$$

Together with (4.3.55) let us consider the stochastic differential equation

$$dY_t = \beta_t(Y_t)\,dt - \sigma_t^k(Y_t)\,dw_t^k - \rho_t^r(Y_t)\,d\hat w_t^r, \quad 0 \le t \le T, \quad Y_0 = y, \tag{4.3.58}$$

where

$$\beta_t(y) = -b_t(y) + \sigma_t^{ik}(y)D_i\sigma_k(t, y) + \rho_t^{ri}(y)D_i\rho_t^r(y), \quad t \in [0, T],\, y \in \mathbb{R}^d.$$

By the Itô-Wentzell formula from [23], for

$$U_t(y) := v_t(Y_t(y))$$

we have (to ease the notation we omit the parameter y from $Y_t(y)$)

$$\begin{aligned} dv_t(Y_t) = &(Lv_t(Y_t) + f_t(Y_t))\,dt + (M^k v_t(Y_t) + g_t^k(Y_t))\,dw_t^k + \mathcal{N}^r v_t(Y_t)\,d\hat w_t^r \\ &+(\beta_t^i D_i v_t(Y_t) + a_t^{ij}D_{ij}v_t(Y_t))\,dt - \sigma_t^{ik}D_i v_t(Y_t)\,dw_t^k - \mathcal{N}^r v(Y_t)\,d\hat w_t^r \\ &- \sigma_t^{ik}D_i(M^k v_t(Y_t) + g^k(Y_t))\,dt - \mathcal{N}^r\mathcal{N}^r v_t(Y_t)\,dt. \end{aligned} \tag{4.3.59}$$

Due to cancellations on the right-hand side of (4.3.59) we obtain

$$
\begin{aligned}
dU_t(y) =& \{\gamma_t(Y_t(y))U_t(y) + \phi_t(Y_t(y))\}\, dt \\
&+ \{\mu^k_t(Y_t(y))U_t(y) + g^k_t(Y_t(y))\}\, dw^k_t, \quad U_0(y) = \psi(y),
\end{aligned}
$$

where

$$
\gamma_t(x) := c_t(x) - \sigma^{ki}_t(x)D_i\mu^k_t(x), \quad \phi_t(x) = f_t(x) - \sigma^{ki}_t(x)D_i g^k_t.
$$

Notice that in the special case when $f = 0$, $g = 0$, $c = 0$, $\mu = 0$ and $\psi(x) = x^i$ for $i \in \{1, ..., d\}$, we get $\tilde{v}^i_t(Y_t(y)) = y^i$ for $i = 1, ..., d$, where $\tilde{v}^i$ is the solution of (4.3.55)-(4.3.56) with $f = c = 0$, $g = \mu = 0$, $\sigma = 0$ and $\psi(x) = x^i$. Hence for each $t \in [0, T]$ the mapping $y \to Y_t(y) \in \mathbb{R}^d$ has an inverse, Y^{-1}_t, for almost every ω, and the mapping $x \to \tilde{v}_t(x) = (\tilde{v}^i_t(x))^d_{i=1}$, defined by the continuous random field $(\tilde{v}^i_t)_{(t,x)\in H_T}$ gives a continuous modification of Y^{-1}_t. Also, we can write $v_t(x) = U_t(Y^{-1}_t)$. The idea of this transformation follows [27] where this was used to show the existence of the inverse of flows given by diffusion processes, and to describe their dynamics.

Set $\bar{U}_t(y) = \bar{v}_t(\bar{Y}_t(y))$, where $\bar{v}_t(x)$ and $\bar{Y}_t(y)$ are defined as $v_t(x)$ and $Y_t(y)$ in (4.3.55)-(4.3.56) and (4.3.58), respectively, with $\bar{\mathfrak{D}}$ and $\bar{\rho}$ in place of $\mathfrak{D}$ and ρ.

Introduce the notation $\mathbf{A}_R = \mathbb{B}_R \cap \mathbb{Q}^{d+1}$. Since u and $\bar{u}$ are continuous in both variables,

$$
\sup_{(t,x)\in\mathbb{B}_{\nu R}} |u_t(x) - \bar{u}_t(x)| = \sup_{(t,x)\in\mathbf{A}_{\nu R}} |u_t(x) - \bar{u}_t(x)| \tag{4.3.60}
$$

Let $\nu' = (1 + \nu)/2$ and define the event

$$
H := \left[\sup_{(t,x)\in\mathbb{B}_{\nu R}} |Y^{-1}_t(x)| > \nu' R\right] \cup \left[\sup_{(t,x)\in\mathbb{B}_{\nu' R}} |Y_t(x)| > R\right].
$$

Then

$$
H^c = \left[Y^{-1}_t(x) \in B_{\nu' R}, \forall (t, x) \in \mathbb{B}_{\nu R}\right] \cap \left[Y_t(x) \in B_R, \forall (t, x) \in \mathbb{B}_{\nu' R}\right],
$$

and thus on H^c

$$
Y_t(x) = \bar{Y}_t(x) \quad \text{for } (t, x) \in \mathbb{B}_{\nu' R},
$$

$$
Y^{-1}_t(x) = \bar{Y}^{-1}_t(x) \quad \text{for } (t, x) \in \mathbb{B}_{\nu R}\ ,
$$

and consequently,

$$v_t(x) = \bar{v}_t(x) \quad \text{for } (t,x) \in \mathbb{B}_{\nu R}.$$

Therefore, by (4.3.57) and (4.3.60), and by Doob's, Hölder's, and the conditional Jensen inequalities,

$$E \sup_{(t,x)\in\mathbb{B}_{\nu R}} |u_t(x) - \bar{u}_t(x)|^q \leq E \sup_{t\in[0,T]\cap\mathbb{Q}} |E(\mathbf{1}_H \sup_{(\tau,x)\in\mathbf{A}_{\nu R}} |v_\tau(x) - \bar{v}_\tau(x)||\mathcal{F}_t)|^q$$

$$\leq \frac{q}{q-1}(P(H))^{1/r} E^{1/r'}(\sup_{(\tau,x)\in H_T} |v_\tau(x) - \bar{v}_\tau(x)|^{qr'}) \tag{4.3.61}$$

$$\leq \frac{2^{q-1}q}{q-1}(P(H))^{1/r} V_T \tag{4.3.62}$$

with

$$V_T := E^{1/r'} \sup_{(\tau,x)\in H_T} |v_\tau(x)|^{qr'} + E^{1/r'} \sup_{(\tau,x)\in H_T} |\bar{v}_\tau(x)|^{qr'},$$

for $r > 1$, $r' = r/(r-1)$, provided $q > 1$. By Theorem 3.1.1

$$V_T \leq N E^{1/r'}(\mathcal{K}^{qr'}_{m,p}(T) + \bar{\mathcal{K}}^{qr'}_{m,p}(T)). \tag{4.3.63}$$

We can estimate $P(H)$ as follows. Clearly,

$$P(H) \leq P(\sup_{(t,x)\in\mathbb{B}_{\nu R}} |Y_t^{-1}(x)| > \nu' R) + P(\sup_{(t,x)\in\mathbb{B}_{\nu' R}} |Y_t(x)| > R) =: J_1 + J_2.$$

For $\hat{Y}_t(x) = Y_t(x) - x$ by (4.3.53) we have

$$J_2 \leq P(\sup_{(t,x)\in\mathbb{B}_{\nu' R}} |\hat{Y}_t(x)| > (1-\nu')R) \leq N R^{d+1/2} e^{-\delta(1-\nu)^2 R^2}.$$

Also, we have

$$J_1 \leq \sum_{l=0}^{\infty} P(\exists (t,x) \in [0,T] \times (B_{2^{l+1}\nu' R} \setminus B_{2^l \nu' R}) : |Y_t(x)| \leq \nu R)$$

$$\leq \sum_{l=0}^{\infty} P(\sup_{(t,x)\in\mathbb{B}_{2^{l+1}\nu' R}} |\hat{Y}_t(x)| \geq (2^l \nu' - \nu)R).$$

Using (4.3.53) again gives

$$J_1 \leq N \sum_{l=0}^{\infty} e^{-\delta(2^l\nu' - \nu)^2 R^2} (2^{l+1}\nu' R)^{d+1} \leq N e^{-\delta R^2}$$

We can conclude that

$$P(H) \leq N e^{-\delta R^2}, \tag{4.3.64}$$

where N and δ are positive constants, depending only on d, K and T.

Combining this with (4.3.62) we can finish the proof. The case $q \in (0, 1]$ follows easily from the usual arguments using Lemma 1.2.3. □

4.4 A fully discrete scheme

We now apply our localization result to present a numerical scheme approximating (4.3.45). We make use of the results of [13] on the rate and acceleration of finite difference approximations, which, together with a time discretization and a truncation - whose error can be estimated using Theorem 4.3.1 - yields a fully implementable scheme.

First we introduce the finite difference approximation of an equation with arbitrary data $\tilde{\mathfrak{D}}$. It is slightly different, and in the main aspects, more general, than the one introduced in Section 4.1, so let us introduce the whole formulation. Let $\Lambda_1 \subset \mathbb{R}^d$ be a finite set, containing the zero vector, satisfying the following natural condition: if a subset $\Lambda' \subset \Lambda_1$ is linearly dependent, then it is linearly dependent over the rationals. Set also $\Lambda_0 = \Lambda_1 \setminus \{0\}$. For $h \neq 0$ define the grid

$$\mathbb{G}_h = \{h \sum_{i=1}^{n} \lambda_i : \lambda_i \in \Lambda_1 \cup -\Lambda_1, n = 1, 2, \ldots\}$$

and for $\lambda \in \Lambda_0 \cup -\Lambda_0$, the finite difference operators

$$\delta_{h,\lambda}\varphi(x) = \frac{1}{h}(\varphi(x + h\lambda) - \varphi(x)), \;\; \delta^h_\lambda = \frac{1}{2}(\delta_{h,\lambda} + \delta_{-h,\lambda}) = \frac{1}{2}(\delta_{h,\lambda} - \delta_{h,-\lambda}),$$

and let both $\delta_{h,0}$ and δ^h_0 stand for the indentity operator. For $h \neq 0$ consider the equation

$$dv_t(x) = (\tilde{L}^h v_t(x) + \tilde{f}_t(x))\, dt + \sum_{k=1}^{\infty} (\tilde{M}^{h,k} v_t(x) + \tilde{g}^k_t(x)) dw^k_t \tag{4.4.65}$$

on $[0, T] \times \mathbb{G}_h$, with initial condition

$$v_0(x) = \tilde{\psi}(x). \tag{4.4.66}$$

Here $\tilde{L}^h$ and $\tilde{M}^{h,k}$ are difference operators approximating the differential operators $\tilde{L}, \tilde{M}^k$:

$$\tilde{L}^h = \sum_{\lambda,\kappa\in\Lambda_1} \mathfrak{a}^{\lambda\kappa}\delta^h_\lambda\delta^h_\kappa + \sum_{\lambda\in\Lambda_0}(\mathfrak{p}^\lambda\delta_{h,\lambda} - \mathfrak{q}^\lambda\delta_{h,-\lambda}), \quad \tilde{M}^{h,k} = \sum_{\lambda\in\lambda_1}\mathfrak{b}^{\lambda,k}\delta^h_\lambda,$$

where the coefficients $\mathfrak{a}, \mathfrak{p}, \mathfrak{q}, \mathfrak{b}$ are related to the data $\tilde{\mathfrak{D}}$ through a compatibility condition, see Assumption 4.4.1 below.

Notice that unless $\tilde{\mathfrak{D}}$ is compactly supported (i.e. each component of it is), equation (4.4.65)-(4.4.66) is still an infinite dimensional system of SDEs. Therefore to make the method practical, we truncate the system. In other words, to get an approximation of the solution of (4.3.45)-(4.3.46), we first take a truncation $\mathfrak{D}^R$ of $\mathfrak{D}$, as described below, and then apply the scheme (4.4.65)-(4.4.66) with $\tilde{\mathfrak{D}} = \mathfrak{D}^R$. First we fix a function $\zeta \in C_0^\infty(\mathbb{R}^d)$ such that $\zeta(x) = 1$ for $|x| \leq 1$ and $\zeta(x) = 0$ for $|x| \geq 1 + \epsilon$ for some $\epsilon > 0$. With the notation any $\phi^{(R)}(x) = \phi(x)\zeta(x/R)$ for any $R > 0$ and function ϕ defined on $\mathbb{R}^d$, define

$$\mathfrak{D}^R = (\psi^{(R)}, (a^{(R)})^{(R)}, b^{(R)}, c^{(R)}, \sigma^{(R)}, \mu^{(R)}, f^{(R)}, g^{(R)})$$

Note that the bounds for $\mathfrak{D}^R$ are uniform for, say, $R \geq 1$, and depend only on the bounds for $\mathfrak{D}$ and the derivatives of ζ.

At this point our approximation is a finite dimensional SDE, and the time-discretization of such equations are well studied. For our purposes the suitable choice is the implicit Euler method. Let $n \geq 1$, $\tau = T/n$. Consider the following approximation of (4.4.65)-(4.4.66):

$$v_i = v_{i-1} + (\tilde{L}^h_{\tau(i-1)}v_i + \tilde{f}_{\tau(i-1)})\tau + \sum_{k=1}^{\infty}(\tilde{M}^{h,k}_{\tau(i-1)}v_{i-1} + \tilde{g}^k_{\tau(i-1)})\xi^k_i \tag{4.4.67}$$

for $i = 1, 2, \ldots, n$, where $\xi^k_i = w^k_{\tau i} - w^k_{\tau(i-1)}$, with initial condition

$$v_0 = \tilde{\psi}. \tag{4.4.68}$$

Remark 4.4.1. The concept of a solution of (4.4.65)-(4.4.66), as a process with values in $l_{2,h}$, that is, the space of functions $\phi : \mathbb{G}_h \to \mathbb{R}$ with finite norm $\|\phi\|^2_{l_{2,h}} = \sum_{x\in\mathbb{G}_h}|\phi(x)|^2$, is straightforward. However, similarly to the point of view in Section 4.1, one can also consider (4.4.65)-(4.4.66) on the whole space, that is, for $(t, x) \in H_T$. In this case we look for a v such that the two sides of the equation coincide almost surely for all t as processes in L_2, and we will refer to such a v to be the L_2-valued solution of (4.4.65)-(4.4.66). The analogous concepts will be used for solutions of (4.4.67)-(4.4.68).

Remark 4.4.2. In many applications, including the Zakai equation for nonlinear filtering, the driving noise is finite dimensional. If this is not the case, one needs another level of approximation, at which the infinite sum in (4.4.67) is replaced by its first m terms. We shall not discuss this here.

Finally, recall the setting of Richardson extrapolation. Let $r \geq 1$, V be the $(r+1) \times (r+1)$ Vandermonde matrix $V^{ij} = (2^{-(i-1)(j-1)})$,

$$(c_0, c_1, \ldots, c_r) = (1, 0, \ldots, 0)V^{-1},$$

and for a parametrized family of random fields $(u^h)_{h>0}$, define

$$v^h = \sum_{i=0}^{r} c_i u^{h/2^i}. \tag{4.4.69}$$

Assumption 4.4.1. For every $i, j = 1, \ldots, d$, $k = 1, \ldots$

$$a^{ij} = \sum_{\lambda,\kappa \in \Lambda_0} \mathfrak{a}^{\lambda\kappa} \lambda^i \kappa^j, \ b^i = \sum_{\lambda \in \Lambda_0} (\mathfrak{a}^{0\lambda} + \mathfrak{a}^{\lambda 0} + \mathfrak{p}^\lambda - \mathfrak{q}^\lambda)\lambda^i, \ c = \mathfrak{a}^{00},$$

$$\sigma^{ik} = \sum_{\lambda \in \Lambda_0} \mathfrak{b}^{\lambda,k} \lambda^i, \mu^k = \mathfrak{b}^{0,k}$$

Assumption 4.4.2. For $P \otimes dt \otimes dx$-almost all $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$

$$\sum_{\lambda,\kappa \in \Lambda_0} (2\mathfrak{a}^{\lambda\kappa} - \mathfrak{b}^{\lambda,k}\mathfrak{b}^{\kappa,k}) z^\lambda z^\kappa \geq 0$$

for all $z \in \mathbb{R}^{\#\{\Lambda_0\}}$.

Assumption 4.4.3. The functions $\mathfrak{a}^{\lambda\kappa}$ and their derivatives in x up to order $\max(m, 2)$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $\lambda, \kappa \in \Lambda_1$. The functions $\mathfrak{b}^\lambda = (\mathfrak{b}^{\lambda r})_{r=1}^\infty$ and their derivatives in x up to order $m+1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued functions, bounded by K, for all $\lambda \in \Lambda_1$.

Assumption 4.4.4. The functions $\mathfrak{a}^{\lambda\kappa}, \mathfrak{b}^\lambda, \mathfrak{p}^\lambda, \mathfrak{q}^\lambda, f, g$ as processes with values in $\mathbb{R}, l_2, \mathbb{R}, \mathbb{R}, W_2^m, W_2^m(l_2)$, respectively, are 1/2-Hölder continuous in t with Hölder constant η, where η is a finite random variable.

Clearly, under Assumption 4.4.1, Assumption 4.4.3 implies Assumption 4.3.2 (a)-(b).

As for the following we confine ourselves to the L_2-scale, without weights, we use the shorthand notation $\|\cdot\|_m = |\cdot|_{W_2^m}$, $\|\cdot\| = \|\cdot\|_0$, and similarly for $\mathcal{K}$. The main result of this section is the following.

Theorem 4.4.1. *Let Assumption 4.3.2 (c), 4.3.3 for $p = 2$, and 4.4.1 through 4.4.4 hold for the data $\mathfrak{D}$ with $m > 2r + 3 + d/2$ for an integer $r \geq 1$. Then if τ is sufficiently small, then for any $R \geq 1, h > 0$, with $\tilde{\mathfrak{D}} = \mathfrak{D}^R$, the system of equations* (4.4.67)-(4.4.68) *has a unique solution $(u_i^{R,h,\tau})_{i=0}^n$, and defining its extrapolation of order r by $(v_i^{R,h,\tau})_{i=0}^n$ as in* (4.4.69)*, we have, for any $r' > 1$, $\nu \in (0,1)$*

$$E \max_{i=0,\dots,n} \max_{x \in \mathbb{G}_h \cap B_{\nu R}} |u_{\tau i}(x) - v_i^{R,h,\tau}(x)|^2$$
$$\leq N(e^{-\delta R^2} + h^{2(r+1)} + \tau) E^{1/r'}(1 + \mathcal{K}_m^{2r'}),$$

where N and δ depends on K, d, T, m, ν, and $E|\eta|^{2r'/(r'-1)}$.

This theorem is a simple consequence of the Theorem 4.3.1, the results of [13], which are summarized below in Theorem 4.4.2, and the error estimate for the time-discretization, established in Theorem 4.4.3 below. This can be seen by simply writing the error as

$$u_{\tau i} - v_i^{R,h,\tau} = (u_{\tau i} - u_{\tau i}^R) + (u_{\tau i}^R - v_{\tau i}^{R,h}) + \sum_{j=0}^{r} c_j (u_{\tau i}^{R,h/2^j} - u_i^{R,h/2^j,\tau}).$$

Theorem 4.4.2. *Let Assumptions 4.3.3 for $p = 2$, $\vartheta = 0$ and 4.4.1 through 4.4.3 hold for $\tilde{\mathfrak{D}}$ with m. Then*

(a) For any $\phi \in W_p^m$ and $|\gamma| \leq m$

$$2(D^\gamma \phi, D^\gamma \tilde{L}^h \phi) + \sum_k \|D^\gamma \tilde{M}^{h,k} \phi\|^2 \leq N \|\phi\|_m^2;$$

(b) There is a unique L_2-valued solution $\tilde{u}^h$ of (4.4.65)-(4.4.66)*, and*

$$E \sup_t \|\tilde{u}_t^h\|_m^q \leq N E \tilde{\mathcal{K}}_m^q;$$

(c) If furthermore $m > 2r + 3 + d/2$, then denoting the solution of (4.3.45)-(4.3.46) *with data $\tilde{\mathfrak{D}}$ by $\tilde{u}$, and the extrapolation of $\tilde{u}^h$ of order r by $\tilde{v}^h$ as in* (4.4.69)*, we have*

$$E \sup_t \max_{x \in \mathbb{G}_h} |\tilde{u}_t(x) - \tilde{v}^h(x)|^q \leq N h^{q(r+1)} E \tilde{\mathcal{K}}_m^q,$$

where N depends on K, d, T, q, and m.

Theorem 4.4.3. *Let Assumptions 4.3.3 with $p = 2$, $\vartheta = 0$, and 4.4.1 through 4.4.3 hold with $m+5$. Then for sufficiently small τ there exists a unique L_2-valued* *solution* $\tilde{u}^{h,\tau}$ *to* (4.4.67)-(4.4.68)*, and for any* $r' > 1$

$$E \max_i \|\tilde{u}^h_{\tau i} - \tilde{u}^{h,\tau}_i\|^2_m \le N\tau(1 + E^{1/r'}\tilde{\mathcal{K}}^{2r'}_{m+5}),$$

where N *depends on* K*,* d*,* T*,* m*, and* $E|\eta|^{2r'/(r'-1)}$*.*

Proof. The solvability of (4.4.67)-(4.4.68) can be seen by induction: $\tilde{u}^{h,\tau}_i$ can be constructed from $\tilde{u}^{h,\tau}_{i-1}$ due to the invertibility of the operator $I - \tau\tilde{L}^h_{\tau(i-1)}$ for sufficiently small τ. For further details we refer to [15], Section 3.2.

Let us fix a multiindex γ with $|\gamma| \le m+1$. Substracting (4.4.67) from (4.4.65), we get that the error $e_i = \tilde{u}^h_{\tau i} - \tilde{u}^{h,\tau}_i$ is a W^m_2-valued $\mathcal{F}_{\tau i}$-measurable random variable, $i = 0, \ldots, n$, and its derivative of order γ is the L_2-valued solution of the equation

$$\begin{aligned} D^\gamma e_i =& D^\gamma e_{i-1} \\ &+ D^\gamma \tilde{L}^h_{\tau(i-1)} e_i \tau + \int_{\tau(i-1)}^{\tau i} D^\gamma F_s \, ds \\ &+ D^\gamma \tilde{M}^{h,k}_{\tau(i-1)} e_{i-1} \xi^k_i + \int_{\tau(i-1)}^{\tau i} D^\gamma G^k_s \, dw^k_s \end{aligned} \tag{4.4.70}$$

for $i = 1, \ldots, n$, with zero initial condition, where with the notations $\kappa_1(t) = \kappa_1^n(t) = \lfloor nt \rfloor / n$ and $\kappa_2(t) = \kappa_2^n(t) = (\lfloor nt \rfloor + 1)/n$,

$$F_t = (\tilde{L}^h_t \tilde{u}^h_t - \tilde{L}^h_{\kappa_1(t)} \tilde{u}^h_{\kappa_2(t)} + \tilde{f}_t - \tilde{f}_{\kappa_1(t)}),$$

$$G^k_t = (\tilde{M}^{h,k}_t \tilde{u}^h_t - \tilde{M}^{h,k}_{\kappa_1(t)} \tilde{u}^h_{\kappa_1(t)} + \tilde{g}^k_t - \tilde{g}^k_{\kappa_1(t)}).$$

Introducing the notations

$$\mathfrak{L}_i = \tilde{L}^h_{\tau i}, \, \mathfrak{M}^k_i = M^{h,k}_{\tau i}, \quad \mathfrak{F}_i = \int_{\tau(i-1)}^{\tau i} D^\gamma F_s \, ds, \quad \mathfrak{G}_i = \int_{\tau(i-1)}^{\tau i} D^\gamma G^k_s \, dw^k_s,$$

$$\mathfrak{K}^2_i = \int_{\tau(i-1)}^{\tau i} \|F_s\|^2_{m+1} + \|G_s\|^2_{m+2} \, ds,$$

we can express the difference

$$\begin{aligned} &\|D^\gamma e_i\|^2 - \|D^\gamma e_{i-1}\|^2 \\ &\quad = 2(D^\gamma e_i, D^\gamma \mathfrak{L}_{i-1} e_i \tau + \mathfrak{F}_i) + 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i + \mathfrak{G}_i) \\ &\quad\quad + 2(D^\gamma e_i - D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i + \mathfrak{G}_i) - \|D^\gamma e_i - D^\gamma e_{i-1}\|^2 \\ &\quad = 2(D^\gamma e_i, D^\gamma \mathfrak{L}_{i-1} i e_i \tau + \mathfrak{F}_i) + 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i + \mathfrak{G}_i) \end{aligned}$$

$$
\begin{aligned}
&+ \|D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i + \mathfrak{G}_i\|^2 - \|D^\gamma \mathfrak{L}_{i-1} e_i \tau + \mathfrak{F}_i\|^2 \\
&\le 2(D^\gamma e_i, D^\gamma \mathfrak{L}_{i-1} e_i \tau + \mathfrak{F}_i) + 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i + \mathfrak{G}_i) \\
&+ \|D^\gamma \mathfrak{M}^k_{i-1} e_{i-1} \xi^k_i\|^2 + 2(D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i, \mathfrak{G}_i) + \|D^\gamma \mathfrak{G}_i\|^2. \qquad (4.4.71)
\end{aligned}
$$

The second term on the right-hand side has 0 expectation. By Itô's isometry and integration by parts, we have

$$
E\|D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i\|^2 = \sum_k \|D^\gamma \mathfrak{M}^k_i e_{i-1}\|^2 \tau,
$$

$$
E(D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i, \mathfrak{G}_i) \le \tau N \|D^\gamma e_{i-1}\|^2 + N E \mathfrak{K}^2_i,
$$

$$
E\|\mathfrak{G}_i\|^2 \le E\mathfrak{K}^2_i.
$$

Recall furthermore from Theorem 4.4.2 (a) that

$$
2(D^\gamma e_i, D^\gamma \mathfrak{L}_i e_i) + \sum_k \|D^\gamma \mathfrak{M}^k e_i\|^2 \le N \|e_i\|^2_{m+1}.
$$

Therefore, by taking expectations and summing up (4.4.71) from 1 to j and for $|\gamma| \le m+1$, keeping in mind that $e_0 = 0$, we get

$$
E\|e_j\|^2_{m+1} \le N \sum_{i=1}^{j} \tau E\|e_i\|^2_{m+1} + E\mathfrak{K}^2_i,
$$

and we can conclude by a simple induction that

$$
\max_j E\|e_j\|^2_{m+1} \le (1 - N\tau)^n \sum_{i=1}^{n} E\mathfrak{K}^2_i \le N \sum_{i=1}^{n} E\mathfrak{K}^2_i. \qquad (4.4.72)
$$

Now let $|\gamma| \le m$ and sum up (4.4.71) from 1 to j without taking expectations. We can use Theorem 4.4.2 as before and obtain

$$
\|D^\gamma e_j\|^2 \le N \sum_{i=1}^{n} (\tau \|e_i\|^2_m + \mathfrak{K}^2_i) + \sum_{i=1}^{j} (m^{(1)}_i + m^{(2)}_i + \hat{m}^{(3)}_i),
$$

where

$$
m^{(1)}_i = 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i + \mathfrak{G}_i),
$$

$$
m^{(2)}_i = \|D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i\|^2 - \tau \sum_k \|D^\gamma \mathfrak{M}^k_i e_{i-1}\|^2,
$$

$$
\hat{m}^{(3)}_i = 2(D^\gamma \mathfrak{M}^k_i e_{i-1} \xi^k_i, \mathfrak{G}_i).
$$

We can write $\hat{m}_i^{(3)} = m_i^{(3)} + \bar{m}_i^{(3)}$ with

$$m_i^{(3)} := 2(D^\gamma \mathfrak{M}_i^k e_{i-1}\xi_i^k, \mathfrak{G}_i) - 2(D^\gamma \mathfrak{M}_i^k e_{i-1}, \int_{\tau(i-1)}^{\tau i} G_s^k ds)$$

$$\bar{m}_i^{(3)} := 2(D^\gamma \mathfrak{M}_i^k e_{i-1}, \int_{\tau(i-1)}^{\tau i} G_s^k ds),$$

and after integration by parts,

$$\sum_{i=1}^{j} |\bar{m}_i^{(3)}| \le N \sum_{i=1}^{n} (\tau \|e_i\|_m^2 + \mathfrak{K}_i^2).$$

Thus,

$$E \max_j \|D^\gamma e_j\|^2 \le N \sum_{i=1}^{n} (\tau E\|e_i\|_m^2 + E\mathfrak{K}_i^2) + E \max_j \sum_{i=1}^{j} (m_i^{(1)} + m_i^{(2)} + m_i^{(3)}).$$

Notice that $m_j^{(l)}$ are martingale differences for $l = 1, 2, 3$, so the second term on the right-hand side is estimated through martingale inequalities. We only detail the contribution of $m^{(2)}$, the other terms can be treated similarly. Remember that $\xi_i = w_{t_i} - w_{t_{i-1}}$ and note that by Itô's formula

$$\xi_i^k \xi_i^l - \mathbf{1}_{k=l}\tau = \int_{\tau(i-1)}^{\tau i} (w_s^k - w_{\kappa_1(s)}^k)\, dw_s^l + \int_{\tau(i-1)}^{\tau i} (w_s^l - w_{\kappa_1(s)}^l)\, dw_s^k,$$

and therefore, by the Burkholder-Gundy-Davis inequality, with the notation $i_s = k_1(s)/\tau$ we have

$$\begin{aligned}
&E \max_j \sum_{i=1}^{j} m_i^{(2)} \\
&\le 2E \sup_{t \le T} \sum_k \int_0^t \sum_l (w_s^l - w_{\kappa_1(s)}^l)(D^\gamma \mathfrak{M}_{i(s)}^l e_{i(s)}, D^\gamma \mathfrak{M}_{i(s)}^k e_{i(s)})\, dw_s^k \\
&\le 6E \left(\int_0^T \sum_k |\sum_l (w_s^l - w_{\kappa_1(s)}^l)(D^\gamma \mathfrak{M}_{i(s)}^l e_{i(s)}, D^\gamma \mathfrak{M}_{i(s)}^k e_{i(s)})|^2\, ds \right)^{\frac{1}{2}} \\
&\le 6E \left(\max_{i \le n} \sum_k \|D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}\|^2 \int_0^T \|\sum_l (w_s^l - w_{\kappa_1(s)}^l) D^\gamma \mathfrak{M}_{i(s)}^l e_{i(s)}\|^2\, ds \right)^{\frac{1}{2}}
\end{aligned}$$

We can continue with estimating the maximum on the right-hand side of the last inequality by taking the sum over i and using using Young's inequality $2ab \le$

$\tau a^2 + b^2/\tau$ with

$$a^2 := \sum_{i\le n}\sum_k \|D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\|^2$$

$$b^2 := \int_0^T \|\sum_l (w^l_s - w^l_{\kappa_1(s)}) D^\gamma \mathfrak{M}^l_{i(s)} e_{i(s)}\|^2\, ds,$$

to get

$$\begin{aligned}
&E\max_j \sum_{i=1}^{j} m_i^2 \\
&\le N\sum_{i=1}^{n}(\tau E\|e_i\|^2_{m+1}) + (1/\tau)\int_0^T E\|\sum_l (w^l_s - w^l_{\kappa_1(s)}) D^\gamma \mathfrak{M}^l_{i(s)} e_{i(s)}\|^2\, ds \\
&\le N\sum_{i=1}^{n}(\tau E\|e_i\|^2_{m+1}) + N\int_0^T E\|e_{i(s)}\|^2_{m+1}\, ds.
\end{aligned}$$

We can conclude that

$$E\max_j \|D^\gamma e_j\|^2 \le N\sum_{i=1}^{n}(\tau E\|e_i\|^2_{m+1} + E\mathfrak{K}_i^2),$$

and upon summing up over $|\gamma| \le m$ and invoking (4.4.72), we get

$$E\max_j \|e_j\|^2_m \le N\sum_{i=1}^{n} E\mathfrak{K}_i^2 \le N\sup_{s\le T}\left(E\|F_s\|_{m+1} + E\|G_s\|_{m+2}\right).$$

To estimate $E\|F_s\|_{m+1}$ notice that due to the 1/2-Hölder continuity of $\tilde{f}$ and the coefficients of $\tilde{L}^h$, we can write

$$E\|F_t\|^2_{m+1} \le N\tau(1 + E^{1/r'}\sup_{s\in[0,T]}\|\tilde{u}^h_s\|^{2r'}_{m+3}) + NE\|\tilde{u}^h_{\kappa_2(t)} - \tilde{u}^h_t\|^2_{m+3}.$$

Furthermore, by the definition of κ_2 and the equation of u,

$$E\|\tilde{u}^h_{\kappa_2(t)} - \tilde{u}^h_t\|^2_{m+3} = E\Big|\Big|\int_t^{\kappa_2(t)} \tilde{L}^h_s u_s + \tilde{f}_s\, ds + \int_t^{\kappa_2(t)} \tilde{M}^{h,k}_s \tilde{u}^h_s + \tilde{g}^k_s\, dw^k_s\Big|\Big|^2_{m+3}$$

$$\le 2E\left(\int_t^{\kappa_2(t)} \|\tilde{L}^h_s\tilde{u}^h_s + \tilde{f}_s\|_{m+3}\, ds\right)^2 + 2E\int_t^{\kappa_2(t)} \|\tilde{M}^{h,k}_s u_s + \tilde{g}^k_s\|^2_{m+3}\, ds,$$

yielding

$$\sup_{t\in[0,T]} E\|F_t\|^2_{m+1}$$

$$\leq \tau N(1 + E^{1/r'} \sup_{s\in[0,T]} \|\tilde{u}_s^h\|_{m+5}^{2r'} + E \sup_{s\in[0,T]} \|\tilde{f}_s\|_{m+3}^2 + E \sup_{s\in[0,T]} \|\tilde{g}_s\|_{m+3}^2).$$

Noticing that

$$\sup_{s\in[0,T]} \|\tilde{f}_s\|_{m+3}^2 \leq \xi^2 T + \int_0^T \|\tilde{f}_s\|_{m+3}^2\, ds,$$

similarly for g, and invoking the estimate from Theorem 4.4.2 (b), we get

$$\sup_{t\in[0,T]} E\|F_t\|_{m+1}^2 \leq \tau N(1 + E^{1/r'} \tilde{\mathcal{K}}_{m+5}^{2r'}).$$

Similarly we can prove that

$$\sup_{t\in[0,T]} E\|G_t\|_{m+2}^2 \leq \tau N(1 + E^{1/r'} \tilde{\mathcal{K}}_{m+5}^{2r'}),$$

finishing the proof. □

Remark 4.4.3. As it can be easily seen from the last step of the proof, Assumption 4.4.4 can be weakened to α-Hölder continuity for any fixed $\alpha > 0$, at the cost of lowering the rate from $1/2$ to $\alpha \wedge (1/2)$.

To decrease the spatial regularity conditions, in particular, the term $d/2$ to d/p, one can use the results of Section 4.1. Under the additional assumptions formulated therein, we have proved the generalizations of the results of [13], and subsequently, of Theorem 4.4.3, to arbitrary Sobolev spaces W_p^m.

Bibliography

[1] Z. Brzezniak, M. Veraar, Is the stochastic parabolicity condition dependent on p and q?, Electron. J. Probab. 17 (2012), no. 56, 1-24

[2] K. Dareiotis, I. Gyöngy, A comparison principle for stochastic integro-differential equations, Potential Anal. 41 (2014), no 4, pp 1203-1222

[3] K. Dareiotis, M. Gerencsér, On the boundedness of solutions of SPDEs, Stoch. Partial Differ. Equ. Anal. Comput. 3 (2015), no. 1, 84-102.

[4] K. Dareiotis, M. Gerencsér, A weak Harnack inequality for stochastic partial differential equations, arXiv:1503.04472

[5] L. Denis and A. Matoussi, L. Stoica, L^p estimates for the uniform norm of solutions of quasilinear SPDE's, Probability Theory and Related Fields , (2005), Volume 133, Issue 4, pp 437-463

[6] E. De Giorgi, Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari, Mem. Accad. Sci. Torino. Cl. Sci. Fis. Math. Nat., 3, 1957, 25-43

[7] M.I. Freidlin, On the factorization of non-negative definite matrices, Theory Probab. Appl. 13 (1968), 354-356

[8] L. Gerencsér, On a class of mixing processes, Stochastics and Stochastic Reports, Vol 26 (1989), 165-191.

[9] M. Gerencsér, I. Gyöngy, Finite difference schemes for stochastic partial differential equations in Sobolev spaces, Appl. Math. Optim. 72 (2015), no. 1, 77-100

[10] M. Gerencsér, I. Gyöngy, Localization errors in solving stochastic partial differential equations in the whole space, arxiv:1508.05535

[11] M. Gerencsér, I. Gyöngy, N.V. Krylov, On the solvability of degenerate stochastic partial differential equations in Sobolev spaces, Stoch. Partial Differ. Equ. Anal. Comput. 3 (2015), no. 1, 52-83.

[12] I. Gyöngy, An introduction to the theory of stochastic partial differential equations, in preparation.

[13] I. Gyöngy, On stochastic finite difference schemes, Stoch. Partial Differ. Equ. Anal. Comput. 2 (2014), no. 4, 539-583

[14] I. Gyöngy and N.V. Krylov, Expansion of solutions of parametrized equations and acceleration of numerical methods, Illinois Journal of Mathematics, 50 (2006), 473-514. Special Volume in Memory of Joseph Doob (1910 - 2004)

[15] I. Gyöngy and A. Millet, On discretization schemes for stochastic evolution equations, Potential Anal. 23 (2005), no. 2, 99-134

[16] N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North-Holland Publishing Company, Amsterdam-Oxford-New York, 1981.

[17] S. N. Kružhkov, A priori bounds and some properties of solutions of elliptic and parabolic equations Mat. Sb. (N.S.), 65(107):4 (1964), 522-570

[18] N.V. Krylov, An analytic approach to SPDEs, Stochastic partial differential equations: six perspectives, Math. Surveys Monogr., 64, Amer. Math. Soc., Providence, RI (1999), 185–242.

[19] N. V. Krylov, A relatively short proof of Itô's formula for SPDEs and its applications, Stoch. Partial Differ. Equ. Anal. Comput. 1 (2013), no. 1, 152-174

[20] N.V. Krylov, Itô's formula for the L_p-norm of stochastic W^1_p-valued processes, Probab. Theory Relat. Fields, 147 (2010), 583–605.

[21] N. V. Krylov, On the Itô-Wentzell formula for distribution-valued processes and related topics. Probab. Theory Related Fields 150 (2011), no. 1-2, 295-319.

[22] N.V. Krylov, *Controlled Diffusion Processes*, Springer-Verlag, Berlin, 1980.

[23] N.V. Krylov, On the Itô-Wentzell formula for distribution-valued processes and related topics, Probab. Theory Relat. Fields 150 (2011), 295–319.

[24] N. V. Krylov, On factorizations of smooth nonnegative matrix-values functions and on smooth functions with values in polyhedra, Appl. Math. Optim. Volume 58, Issue 3 (2008), 373-392.

[25] N. V. Krylov, B. L. Rosovskii, Stochastic evolution equations. J. Soviet Mathematics, 16, 12331277 (1981)

[26] N. V. Krylov, B. L. Rozovskii, : Characteristics of degenerating second-order parabolic Itô equations. J. Soviet Maths. 32, 336348 (1986)

[27] N.V. Krylov and B.L. Rozovskii, Stochastic partial differential equations and diffusion processes, Russian Math. Surveys Vol. 37, 6 (1982), 81–105

[28] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, 1997.

[29] O. A. Ladyženskaja, V. A. Solonnikov, N. N. Ural'ceva, Linear and quasi-linear equations of parabolic type. (Russian) Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I. 1968 xi+648 pp

[30] S. V. Lototsky and J. Zhong, Stochastic evolution systems with constant coefficients. Stoch. Partial Differ. Equ. Anal. Comput., 1 (2013), no. 4, 687-711

[31] J. Moser, On Harnack's theorem for elliptic differential equations. Comm. Pure Appl. Math. 14 (1961), 577-591

[32] J. Nash, Continuity of solutions of parabolic and elliptic equations. Amer. J. Math. 80, 1958, 931954

[33] O. A. Olejnik and E. V. Radkevich, Second Order Equations with Nonnegative Characteristic Form, AMS, Providence 1973

[34] E. Pardoux, Equations aux dérivées partielles stochastiques nonlinéares monotones. Étude de solutions fortes de type Itô, Thèse Doct. Sci. Math. Univ. Paris Sud (1975)

[35] D. Revuz and M. Yor, Continuous martingales and Brownian motion. Third edition. Grundlehren der Mathematischen Wissenschaften, 293. Springer-Verlag, Berlin, (1999)

[36] L.F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations, Philos.Trans. Roy. Soc. London, Ser.A, 210 (1910), 307-357.

[37] H. Triebel, *Interpolation Theory - Function Spaces - Differential Operators*, North Holland Publishing Company, Amsterdam-New York-Oxford, 1978.